DOCUMENT RESUME

ED 196 789 SO 013 075

AUTHOR	Baker, Paul J., Ed.: Wilson, Everett K., Ed.
TITLE	Knowledge Available and Knowledge Needed to Improve Instruction in Sociology.
INSTITUTION	American Sociological Association, Washington, D.C.
SPONS AGENCY	Fund for the Improvement of Postsecondary Education (DHEW), Washington, D.C.: Lilly Endowment, Inc., Indianapolis, Ind.
PUB DATE	Aug 79
NOTE	346p.: Proceedings of a Conference on Knowledge Available and Knowledge Needed to Improve Instruction in Sociology (Washington, DC, 1979).
AVAILABLE FROM	ASA Teaching Resources Center, American Sociological Association, 1722 N Street, N.W., Washington, DC 20036 ($7.00).
EDRS PRICE	MF01/PC14 Plus Postage.
DESCRIPTORS	*Educational Assessment: *Educational Needs: *Educational Objectives: Faculty Promotion: Grading: Higher Education: *Instructional Improvement: *Sociology: Student Characteristics; Student Evaluation of Teacher Performance; *Teaching Methods

ABSTRACT

This document contains 20 papers which were delivered at a conference on college-level sociology teaching. The aim of the conference was to provide a forum where sociologists could exchange ideas on how to improve sociology instruction. Thirty-four sociologists from state universities and private and city colleges participated in the conference. The document is presented in five major sections--each of which corresponds to a major conference theme. Section I, Rationale, offers an inquiry into the teacher/learning process and suggests that sociology teaching will be more credible if sociology teachers make more use of the classroom as a laboratory for testing central tenets of scientific inquiry. Section II, Planning, offers nine papers on various aspects of assessing objectives, teaching methods and curricula, and students. Section III, Performance, includes two papers on analyzing the classroom environment and assessing student participation in sociology-related field experiences. Section IV, Appraisal, contains six papers discussing grading, assessing critical thinking, and evaluating teacher effectiveness and course innovations. The final section, Overview, offers two papers which discuss the bureaucratic and disciplinary contexts of sociology teaching and the conflict over whether teachers should be rewarded for publishing or good teaching. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED196789

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Dr. Russell
R. Dynes

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

KNOWLEDGE AVAILABLE AND KNOWLEDGE NEEDED TO IMPROVE INSTRUCTION IN SOCIOLOGY

PROCEEDINGS OF A WORKSHOP

KAKN

Edited by:

PAUL J. BAKER

EVERETT K. WILSON

August, 1979

S0 013 075

THE ASA PROJECTS ON TEACHING UNDERGRADUATE SOCIOLOGY

The ASA Projects on Teaching are developing resources to assist sociologists in their teaching role. More than 150 sociologists from four year, two year and university departments work on aspects of (a) curriculum, (b) teacher development—for faculty and graduate students, and (c) institutional contexts in which teaching is carried out. In existence since the Fall of 1974, ASA Projects on Teaching have sponsored teaching workshops, worked cooperatively with state and regional sociological and social science associations (assisting on meeting programs and providing speakers and panelists) and been represented at the annual ASA meeting. A network of resource specialists in teaching is being developed to assist departments and other units of the profession requesting services. Through the ASA Teaching Resources Center the Projects make available to interested sociologists materials on teaching—many of which were prepared by units of the Project. The Projects on Teaching are official programs of the American Sociological Association and are assisted by grants from the Fund for the Improvement of Postsecondary Education (HEW) and the Lilly Endowment, Inc. (Indianapolis). The ASA issues the bi-monthly ASA Teaching Newsletter. The Project has its roots in efforts of the ASA Section on Undergraduate Education. The coordinator of the ASA Projects is Dr. Hans O. Mauksch. For information about subscriptions to the ASA Teaching Newsletter or a list of documents available through the ASA Teaching Resources Center, write to:

Dr. Lawrence J. Rhoades
ASA Teaching Resources Center
American Sociological Association
1722 N Street, N.W.
Washington, D.C. 20036

Sociologists interested in learning more about the ASA Projects and its resources and work are invited to write to:

Dr. Hans O. Mauksch
ASA Projects on Teaching Undergraduate Sociology
Section of Behavioral Sciences
TD3-W Medical Center
University of Missouri
Columbia, MO 65201

KNOWLEDGE AVAILABLE AND KNOWLEDGE NEEDED TO IMPROVE INSTRUCTION IN SOCIOLOGY (KAKN)

The KAKN effort--the workshop, its planning and follow-up--was sponsored by the ASA Projects on Teaching Undergraduate Sociology supported by grants from the Lilly Endowment, Inc. and the Fund for the Improvement of Postsecondary Education (HEW).

THE ASA PROJECTS ON TEACHING UNDERGRADUATE SOCIOLOGY

The ASA Projects on Teaching are developing resources to assist sociologists in their teaching role. More than 150 sociologists from four year, two year and university departments work on aspects of (a) curriculum, (b) teacher development--for faculty and graduate students, and (c) institutional contexts in which teaching is carried out. In existence since the Fall of 1974, ASA Projects on Teaching have sponsored teaching workshops, worked cooperatively with state and regional sociological and social science associations (assisting on meeting programs and providing speakers and panelists) and been represented at the annual ASA meeting. A network of resource specialists in teaching is being developed to assist departments and other units of the profession requesting services. Through the ASA Teaching Resources Center the Projects make available to interested sociologists materials on teaching--many of which were prepared by units of the Project. The Projects on Teaching are official programs of the American Sociological Association and are assisted by grants from the Fund for the Improvement of Postsecondary Education (HEW) and the Lilly Endowment, Inc. (Indianapolis). The ASA issues the bi-monthly ASA Teaching Newsletter. The Project has its roots in efforts of the ASA Section on Undergraduate Education. The coordinator of the ASA Projects is Dr. Hans O. Mauksch. For information about subscriptions to the ASA Teaching Newsletter or a list of documents available through the ASA Teaching Resources Center, write to: Dr. Lawrence J. Rhoades, ASA Teaching Resources Center, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Sociologists interested in learning more about the ASA Projects and its resources and work are invited to write to: Dr. Hans O. Mauksch, ASA Projects on Teaching Undergraduate Sociology, Section of Behavioral Sciences, TD-3 Medical Center, University of Missouri, Columbia, MO 65201.

Conference Participants

Paul J. Baker
Illinois State University

Michael Bassis
University of Rhode Island

Hubert M. Blalock
University of Washington

Ted K. Bradshaw
University of California/Berkeley

Gail Brady
University of Missouri

Frederick Campbell
University of Washington

William D'Antonio
University of Connecticut

James Davis
Harvard University

Richard Dukes
University of Colorado

Bill Ewens
Michigan State University

William Gamson
University of Michigan

Zelda Gamson
University of Michigan

Richard Gelles
University of Rhode Island

Charles A. Goldsmid
Oberlin College

Max Heirich
University of Michigan

Carla Howery
University of Wisconsin/Milwaukee

Ruth Jacobs
Boston University

David Karp
Boston College

Wendy Looman
ASA Teaching Project

Michael Malec
Boston College

Gerald Marwell
University of Wisconsin

Hans O. Mauksch
University of Missouri

Reece McGee
Purdue University

Sharon McPherron
St. Louis Community College

Lawrence J. Rhoades
ASA

Thomas Rice
Denison University

John Schnabel
West Virginia University

Daniel Solomon
U.S. Bureau of the Census

Robert Stauffer
Kalamazoo College

Charlotte Vaughan
Cornell College

Theodore C. Wagenaar
Miami University

Stephen Wieting
University of Iowa

Everett K. Wilson
University of North Carolina
Chapel Hill

Jim Michaels
VPI and State University

Preface

The ASA Projects on Teaching Undergraduate Sociology have generated scores of activities to improve sociology instruction. The Projects have sponsored workshops, task groups, consulting specialists, and have collaborated with those state and regional societies that have put effective teaching high on their agenda. For the first time, many sociology instructors are enjoying the support, encouragement, and stimulation of colleagues across the country. At the same time, such activities have underscored a critical need, the assimilation and assessment of the knowledge, theoretical and empirical, that bears on the role of the professor of sociology.

It was this need that Paul Baker expressed, about three years ago, to Charles Goldsmid: the need to study the intellectual underpinnings of sociology instruction. Goldsmid, Director of the ASA Teacher Development Project, confirmed that others involved in that project held like views. He noted the special interest of Hans Mauksch, Director of the ASA Projects on Teaching. Goldsmid suggested that we organize a conference on Knowledge Available and Knowledge Needed (KAKN) to Improve Instruction in Sociology, promising needed resources and support. Along with Goldsmid, Hans Mauksch and Ted Wagenaar agreed to serve as members of a planning committee.

Three early decisions by this committee provided guidelines. First, we sought participation from a broad cross section of sociologists--people from different types of schools, differing in rank and age and in theoretical and methodological inclinations. (This decision to maximize diversity meant that we would forego any easy--much less, profound--consensus emerging from a sixty-hour conference.) Second, the Planning Committee did not promote any single pedagogical model. The KAKN Conference was designed for open inquiry, not quick closure. Third, we planned an open process with several stages in preparing for the Conference: constructing and distributing a questionnaire to gather colleagues' views, forming working committees based on interests expressed in the questionnaire, permitting each working committee to develop its own agenda and division of labor. In short, we did not envision a panel of experts who would present definitive answers; rather, we invited people to begin a process of inquiry which remains, today, an ongoing enterprise. These three decisions rest on our conviction that the ASA needs a network of teachers and scholars who can learn from each other without imposing a particular brand of sociological or pedagogical dogma.

Table of Contents

Page
List of Conference Participants i
Preface ii
Chapter

I. RATIONALE FOR KAKN CONFERENCE

1. Inquiry Into the Teacher/Learning Process: Trickery, Folklore, or Science? 1
Paul J. Baker

II. PLANNING

A. Assessing the Ends We Serve

2. Assessing the Ends We Serve: We Know Where We're Going Don't We? 10
Michael S. Bassis
3. Sociology and Liberal Education: Some Observations 16
Robert E. Stauffer
4. Applied Sociology and Community Organizations: Building Links for Learning and Working 32
R. Perrucci, H. R. Potter, C. R. Huff
5. Identifying Course Goals: Domains and Levels of Learning 45
Charlotte A. Vaughan
6. Why Formalize the Aims of Instruction? 61
Charles A. Goldsmid

B. Assessing the Means We Use

7. Assessing the Means We Use: Towards a Rational Model of Course Development 80
John F. Schnabel, Richard J. Gelles, William D'Antonio

C. Assessing the People We Teach

8. Selective Recruitment for Successful Sociology Instruction 105
Everett K. Wilson
9. Assessing the People We Teach: Aids to Course Planning 125
Max Heirich
10. Finding Optimal Instructional Programs and Procedures for Individual Students 145
Daniel Solomon

Table of Contents Continues

Chapter Page

III. PERFORMANCE

D. Assessing Performance in the Classroom

11. Learning the Sociology of the Classroom 168
Thomas J. Rice, Ruth Harriet Jacobs, David A. Karp

E. Assessing Performance Beyond the Classroom

12. Letter to FUDGE 185
James A. Davis, Richard Dukes, William A. Gamson

IV. APPRAISAL

F. Assessing Learning Outcomes for Students

13. Grading and the Process of Developing Unfettered, Self-Governing Persons 194
Bill Ewens
14. Learning Sociology and Assessing Critical Thinking 210
Paul J. Baker
15. Designs and Procedures for Evaluating Course Innovations 248
James W. Michaels

G. Assessing the Effectiveness of Teachers

16. Assessing Teacher Effectiveness 259
Theodore C. Wagenaar
17. Teacher Effectiveness: Some Preliminary Questions 273
Michael A. Malec
18. Assessing Teacher Effectiveness in Higher Education: Why, When, Where, and Perhaps Even How? 284
Gerald Marwell

V. OVERVIEW

H. Assessing the Organizational Context of the Teaching - Learning Process

19. The Bureaucratic and Disciplinary Contexts of Teaching 296
Gail Brady, Ted K. Bradshaw, Zelda Gamson,
Hans O. Mauksch, Reece McGee, Lawrence J. Rhoades
20. Increasing the Reward for Teaching 309
Reece McGee

During the fall of 1977, the Planning Committee identified thirty-five colleagues with a clear commitment to the study of intellectual issues germane to sociology instruction. These people were invited to come and work at a three day conference to be held in April of 1978. The invitation outlined the workshop's goals.

> The profession of sociology needs to articulate a basic agenda of key research problems that must be faced if the craft and science of teaching sociology are to be cumulative. . . . We need also to identify, review, and synthesize what is now known. But there are basic conceptual and theoretical concerns as well as problems of methodology and research design. These issues are crucial in creating a more intellectually stimulating climate for teachers of sociology. In short, the proposed workshop is needed to [identify, develop and advance] basic knowledge on the teaching of sociology.

The invitation also included an open-ended set of questions asking: (1) What is reliably known about the teaching of sociology? (2) What sorts of knowledge are lacking? (3) What types of knowledge are most needed in order to improve the teaching of sociology? (4) What sociological knowledge, empirical or theoretical, is most applicable to a better understanding of instruction in sociology? (5) What realms of inquiry, pertinent to sociology instruction, could various persons probe and assess as contributions to the workshop? The returned questionnaires became prime working documents for planning the KAKN Conference. Using these returns, we so prepared the agenda as to make best use of the interests and talents of participants.

The scheme we adopted for the Conference agenda stresses process: planning (preparation for teaching a course), performance, and appraisal or evaluation. These three stages in the process of teaching and learning were subdivided into eight points of inquiry. Committees were then formed for each of the eight problems, and a coordinator was asked to develop an agenda compatible with the special interests of each committee member. These eight issues provide the Table of Contents for this report.

Each committee was charged with two tasks. The first was to identify what is reliably known; the second, what most needs knowing. Tentative answers were to stem from a review of the literature on higher education and scholarly work on teaching, especially sociology teaching and especially from the discipline itself. We provided each committee with a brief commentary on issues deserving attention, as follows.

I. Planning

1. Assessing the Ends We Serve. What is it that sociology teachers really want to accomplish? Do they wish to convey a perspective? information? skills? methods of inquiry? thinking processes? Across all courses? In particular courses? With what types of students? How do we determine the goals of instruction for different courses? What resources are available to teachers of sociology as they (and their departments) try to specify what they are to teach (or what they hope their students will learn)?

2. Assessing the Means We Use. What resources are available to faculty as they plan their courses? What resources are available to chairpersons who are coordinating the process of planning the teaching program? What do we know about various techniques, devices, and strategies which our colleagues have used with some success? What are the range and depth of our knowledge about such resources? Which means are most appropriate for which goals of instruction? How can various means and materials best be classified for other interested teachers?

3. Assessing the People We Teach. As teachers plan their courses, how do they take account of the characteristics of their students? What is the relationship between the characteristics of those recruited (students) and the products of the classroom, i.e., the learning? What can we learn from theory and research about the fit among student characteristics, sociological content, and instructional situation?

II. Performance

4. Assessing Performance in the Classroom. What goes on in the classroom? How do teachers and students define the classroom as a social space? How do these definitions of the classroom influence what does and can go on in a 50-minute time period? What do we know about the learning outcomes of two conventional classroom teaching methods: lectures and discussions? What does the literature from small groups research contribute to our knowledge of the classroom?

5. Assessing Performances Beyond the Classroom. Since the turn of the century, sociology teachers have seen the need to extend learning experiences beyond the conventional classroom operations, as in field work and internships. And with constant technical developments, new instructional possibilities emerge--simulations and games, videotaped material for analysis, computer-assisted instruction (CAI), compressed time films of events, and the like. The central question is this: What kinds of instructional techniques produce what kinds of learning under what conditions? And for what types of students? What do we know about the learning gains from CAI and related efforts to involve students in research? (It will probably be important for this group, as for some others, to specify types of learnings --cognitive, affective, and the like.)

III. Appraisal

6. Assessing Learning Outcomes for Students. A major impediment to the study of teaching as a serious, intellectual enterprise is the inability or failure to measure adequately the crucial dependent variable, student learning. What do we really know about the reliability and validity of existing evaluation instruments and procedures which assess the learning claims made by teachers? How does the adequacy of measures of learning vary with stages of the learning process (see Mead, Bloom, Piaget)? What do we know about grading systems?

7. Assessing the Effectiveness of Teachers. What criteria do we use, what standards do we apply to measure effective teaching? In the avalanche of literature on evaluating teaching, what do we know, surely, and what do we yet need to know about the evaluation of sociology instruction? What do we know about the impact of evaluation on the teacher? on students? on administrators? about the politics of evaluation? Some believe that standardized instruments developed by such agencies as ETS (Educational Testing Service), Purdue, and Kansas State University offer adequate evaluation systems. Is this so, when judged by the conventions of sociological scholarship? What do we know about other means of assessing the effectiveness of instruction in sociology?

IV. Overview

8. Assessing the Organizational Context of the Teaching-Learning Process. How is the process of planning, performing, and appraising carried out at different types of schools? Recognizing that different sorts of

schools embody different types of constraints and opportunities, what sorts of innovations emerge across this spectrum? How do the need and the process of innovation vary across different types of schools? What do we know about the successes and failures of efforts at planned change in the teaching-learning process? What reliable knowledge do we have about organizational reward systems? Can we assess current trends in higher education with any hope of anticipating the working environment of sociology instructors in the 1980s?

Such queries as these were posed in the hope of clarifying our task. Actual procedures were altogether fluid, each group being free to set its own agenda, appropriate to the interests already expressed by its members. Procedures varied by group. In some, each member wrote a position paper. In others, members prepared preliminary documents which were later synthesized in a single statement. Each group was free to begin its inquiry with open minds as to the products for which various members would be responsible. But all participants brought some sort of working document to the KAKN Conference. During those two and one-half days, much time was spent by each of the eight groups, in working sessions where questions were sharpened and ongoing investigations were thoughtfully scrutinized.

The KAKN Conference was designed to generate a set of working papers which would be further developed in following months. The Conference also had a second purpose, the development of a network of teacher-scholars who can learn from each other and arouse the interest of a wider circle. This second goal is ambitious and risky. For research on teaching is notoriously difficult; even the most skilled specialists strike out all too often. How then can the sociologist who is chiefly a teacher hope to do research, or synthesize the fruits of research with the skill to warrant fellow sociologists' approval? If the KAKN Conference was intended to strengthen teaching and the prestige of teachers and if, then, the products are not well received, or do not lead to fruitful research, the whole thing could backfire.

But this, of course, is to misconstrue the aim of the Conference. It was not conceived as the end, but as a needed beginning. We did not expect these products to be pivotal intellectual accomplishments. Rather, they are good starters for further efforts to be pursued in various settings. They are benchmarks for further work. They are part of a process to be kept in motion by a network of sociology instructors who are examining their capacity to nurture learning.

Doubtless we are concerned with the perplexing problem of strengthening the prestige of the sociology teacher. A more sophisticated knowledge base would doubtless enhance that prestige. But perhaps the most significant obstacle to that end is the lowly status of the client. But our concern, really, is not so much with prestige as it is with professional integrity. As professors of sociology who claim to be experts on such matters as groups, human interaction, and socialization, we are practicing very little of what we profess. Is our teaching in accord with the accumulated scientific knowledge we ask our students to learn? Of course one way out of the "What we say, what we do" dilemma is to assert that sociologists have no reliable knowledge of practical worth; hence the use of common sense judgments in the classroom is justified. If, on the contrary, we make high claims for sociological knowledge, we are impelled to support those claims in the test of practical application. And since teachers have great freedom in the classroom, it seems reasonable to review such knowledge as we have in the service of changes likely to improve sociology instruction. Some members of our research guild may demur at some of the conclusions in these chapters. But we are hopeful that many teachers of sociology will see, in reading them, new opportunities to join their scholarly interests with their instructional practices. Such thoughtful integration by college teachers is the first step away from common sense hunches and show-and-tell gimmickry.

Many of the KAKN chapters make use of what is known about the teaching of sociology (indeed there are extensive and valuable resources for teachers of sociology). In almost no areas of instruction in our field, need a sociologist or student of instruction start without benefit of prior work. While much of this literature is descriptive, some, and especially more recent work provides empirical tests or systematic applications of theory to instruction. The scope of resources available is suggested by the third edition of the ASA Projects' Teaching Sociology: An Annotated Bibliography (August, 1979) which contains 470 citations to works specifically on the teaching of sociology.

We write about these Conference chapters in academic terms; but we are mindful that academic products require hard and persistent organizational work. Many sociologists have devoted long hours to bring about the KAKN Conference. We are grateful to the coordinators of the working groups:

Michael Bassis, John Schnabel, Thomas Rice, William Gamson, Ted Wagenaar, and Zelda Gamson. They provided useful guidance at critical stages in the Conference work. Hans Mauksch and Ted Wagenaar helped on many critical questions as we planned the work of colleagues scattered from coast to coast. We relied, regularly, on the skilled services of Wendy Looman who made the arrangements for the KAKN Conference in Southfield, Michigan and kept all participants informed. Sharon Baker typed and edited the manuscript with speed, precision, and a good ear for detecting gobbledygook. More than once she helped unscramble an unintelligible sentence by sociologists whose language failed to match their profundity. To all, we are grateful--but especially to Charles Goldsmid for his sound management, wise counsel, energy, enthusiasm, and support. He is a good sociologist, a good colleague, a good Director of the ASA Project on Teacher Development, and a good steward of the funds that supported the KAKN Conference.

Paul J. Baker, Illinois State University
Everett K. Wilson, University of North
Carolina - Chapel Hill

Inquiry Into the Teaching/Learning Process: Trickery, Folklore, or Science?

Paul J. Baker

Illinois State University

Introductory Comments

During the past four years, I have served as consultant, reviewing and evaluating a number of conventional and innovative programs of college teaching. My observations have brought me to one basic generalization: the prevailing style of college teaching is an unending monologue in which learned professors pour millions of words upon unlearned students. I often see this monologue in metaphorical terms. Visualize a train station where students and faculty are milling about. Eventually, students board the train, while faculty remain on the platform. The train pulls away from the station and the professors run along the tracks. Every two or three days--depending on schedules--students raise the windows and the professors shout a few chosen words; then the windows are lowered as the train keeps moving. At the end of eighteen weeks, a barrier abruptly halts the professors while the train continues to move down the tracks. Watching the train from a distance, the professors notice the windows being raised one final time; the students are throwing out their textbooks.

Some of my colleagues object to this metaphor. It seems too cynical and places a teacher in the awkward role of outsider. I do not want to place undue value on the metaphor, but hope it serves the purpose of helping us re-examine our ideas about the process of teaching and learning. Going one step beyond the train metaphor, I want to suggest that questions of teaching and learning can be explored fruitfully by recognizing three sources of knowledge: trickery, folklore, and science. This classification is a heuristic device intended to stimulate new ways of thinking about teaching and learning. They are not definitive models of sagacity.

Trickery

A few years ago, I conducted a survey of undergraduate teaching programs in Big Ten Universities. With intended profundity, a chairperson of a distinguished department summarized his philosophy of education: "You realize, of course, that teaching is nothing but a bag of tricks." This statement represents the sentiments of many sociologists who proudly sustain maximum distance from any rigorous assessment of undergraduate learning experiences.

Teaching in this context seems to be a medicine show in which some professors enjoy success with gimmicks, while others appear to be inept.

Tricks of the teaching trade come in two varieties. The first form is human relations gimmicks in which the wonders of T-Groups, behavioral modification, encounter groups, and values clarification have the apparent capacity to wash away the tedium of classroom teaching. The tricks of psychology have become a growth industry in recent years. While no claims of deification are made, the human relations style has all the trappings of magical humanism. The second type of trickery comes from the world of electronic technology. On several occasions, inventors and foundation executives have announced the coming educational revolution in which television, learning machines, computers, and other gadgetry will replace the obsolete classroom teacher. These grand predictions have been reassessed soberly in recent times. Much discussion still revolves around the advantages of films, TV, CAI, and multi-media approaches, but few people are willing to proclaim that a technical panacea lies just over the horizon. The tricks of technology are fascinating, but perplexing problems still remain as to how educators are going to integrate hardware, software, and teacher roles.

I do not want to dismiss as worthless the tricks of teaching. Some thoughtful people have devoted careful attention to gimmicks which work for them. An assessment of the tricks--both humanistic and technical--fails however to disclose a coherent sense of the teaching/learning process. The question must be asked: can the development of teaching foster more than a series of mindless gimmicks?

Folklore

Folklore refers to the vast body of common sense ideas, the taken-for-granted beliefs and behaviors which often seem right, or convincing over coffee or in heated debate with teenage children. These are the views that the sociologist constantly attacks in the classroom. Assuming the role of scientist, he or she enjoys informing credulous students that common sense is invariably false because oversimplified, it fails to state the conditions under which verification occurs. Yet, at the same time, the sociologist as teacher is supported by a thick layer of folklore which sustains its practitioners. Let me mention a few of the myths which are cherished by many sociology teachers.

Perhaps the most strongly held belief is the myth that teaching is an ascribed role while research is an achieved role. The ascriptive character of teaching is found in such well-worn phrases as "born teachers," "a natural in the classroom," and "it's impossible to measure good and bad teaching." Research activities, on the other hand, supposedly require specialized training under recognized experts. Researchers feel confident about their judgments of scholarly achievement. The length of a vita and publication citations quickly establish a clear index of success or failure. Since the efficacy of teaching is hard to measure, it is often assumed that all teachers teach about the same. This commonly accepted definition of the situation often places teachers on the defensive; for how claim distinctive credit for what others do equally well? The hard working, creative teacher's light is hid under a bushel--and outshone, in any case, by the glaring light of scholarly recognition.

A second custom firmly entrenched in our priestly tradition can best be described as the trickle-down myth. Proper deference is paid to the laying on of hands, in which many sociology classes intone the sacred names of the masters of sociological thought: Marx, Weber, Durkheim, Simmel, Mead, Park. Photographed dutifully in introductory books, these great-greats and their life works are subjected to two-paragraph summaries. The classroom teacher speaks of near-greats, the people who have directed the teacher's dissertation or who once spoke with the teacher at an ASA convention. Next in the hierarchy are the almost-greats. These are the classroom teachers themselves. They have not made it to the top yet, but there is still hope. If they persist in their efforts to publish, they may someday arrive in the inner privileged circle. And finally, there are the soon-to-be-touched, the students who receive a smattering of the great ideas passed down through the academy. Illustrations of the trickle-down myth are abundant: Tönnies is worth remembering for Gemeinschaft and Gesellschaft; Durkheim provides undergraduates no end of confusion over mechanical and organic solidarity; and, Burgess unlocks the secret of Chicago's growth by concentric zones. These fragments of great ideas are somehow worth a few moments of cerebral attention by bored sophomores. Why? Custom has dictated the sacred order of the academy.

A third aspect of the folklore of college instruction concerns the ideal audience worthy of inclusion. Florian Znaniecki, in The Social Role of the Man of Knowledge, reminds us that persons of knowledge must be understood in the context of their audiences. There is a reciprocal bond between the person of knowledge and the audience. An overview of instruction in sociology suggests that there are three basic types of student audiences: (1) professionals in training (graduate students), (2) pre-professionals preparing for training (sociology majors planning on graduate education), (3) non-professionals with no interest in specialized training (students for whom interest in sociology often ends with the first course final exam). To my knowledge, there are no accurate statistics on the number of students in each category. I would like to offer the following estimates: (1) graduate students: one to three percent, (2) majors: five to eight percent, and (3) those students who could care less about the professional squabbles of sociologists: approximately ninety percent. Esoteric debates of academic sociologists are often the substance of the undergraduate curriculum, e.g., Davis/Morre vs. Tumin, Garfinkle vs. Coleman, Lundberg vs. Nisbet, Homans vs. Parsons, or Blummer vs. Lazersfeld. Nursing students, education majors, and aspiring businessmen have little use for the arguments against psychological reductionism, disputes between conflict theorists and functionalists, or the endless feuds among symbolic interactionists. These are the issues of graduate training which prepare one for the subtle battles of academic combat. Yet many undergraduate programs are designed as pale imitations of graduate education. The edifice constructed for one to three percent of the student audience often casts a long shadow over the reading assignments and classroom lectures of the ninety-plus percent. This imbalance suggests that most curricular planning is designed to meet the needs of college faculty who nervously seek approval from the near greats. Little attention is devoted to the educational needs of the vast majority of undergraduate students who encounter academic sociology.

Our list of foibles from the folklore of college teaching has infinite potential. One final example concerns the undisputed moral and intellectual authority of the teacher in the classroom. What teachers say is good for students must indeed be good for students--"For your benefit, you will memorize Durkheim's three types of suicide." Totalitarian power in the classroom creates truth by fiat. One senses this awesome authority in a statement by the Protestant theologican, Paul Tillich: "our tendency as teachers is to throw answers like rocks at the heads of those who have not yet asked the questions" (see Brown, 1971: 15).

The peculiar capacity of one-sided power to distort learning opportunities is illustrated by an intensive study of social science testimony in the 1974 Bradley vs. Millikin desegregation case in Detroit. In her article, "Courtrooms and Classrooms," Eleanor Wolf discusses the unwillingness of social scientists to honestly educate Judge Roth on important points of interpreting research findings regarding desegregation. She writes (1977: 433):

> there was a large amount of agreeing with the judge, fulsome flattery, in addition to the expected deference. Sometimes people praised him for saying something when in fact he had said the opposite. For a student to have such power over his teachers is not helpful to learning . . .

It seems to me that the opposite holds true for classrooms: "for a teacher to have such power over his students is not helpful to learning." The gamesmanship of flattery described by Wolf about the courtroom is often reversed in the classroom where students pretend to know and appear interested while the teacher is left in the dark on the central issue of learning.

Science

I have only scratched the surface of the folklore about college teaching. The reader can extend this list of examples. I turn now, to the question of science as an alternate perspective. During the past three years, several sociologists associated with the ASA Projects on Teaching have spent considerable money on workshops and conferences sharing the "how to's" of teaching. These efforts often focused on improving the tricks of the trade and cultivating a new sense of professional excitement about college teaching. Workshop participants often expressed gratitude; "I really feel good about this workshop. For the first time in my career, the ASA seems interested in me as a teacher." Much attention has been devoted to improving gimmicks and generating a counter-culture which places teaching on a parity with research. This effort has been of great value to many teachers of sociology; coming out of the closet, we have gained a new sense of professional liberation.

Participants in the ASA Projects on Teaching often refer to their collective efforts as a social movement. But the image may be inflated, the gains shallow and temporary if teachers of sociology fail to practice what they teach. Authentic professional recognition will not be bestowed on teachers of sociology until they generate a scholarly tradition about teaching and learning. Teachers who make their living as experts on such topics as small groups, self-identity, human interaction, social roles, and education must surely recognize the irony of their situation: failure to generate a body of scholarly literature about the social scene--their own classrooms--with which they are most familiar.

To stimulate scholarly activity on behalf of teaching and learning raises many questions which cannot be answered quickly. In the 1930s, Howard P. Becker and Harry E. Barnes wrote a history of sociology entitled Social Thought from Lore to Science. Can the teaching of sociology make the same transition from lore to science? More recently, Nicholas Mullins suggested that scholarly developments in sociology are often located in theory groups. In the spirit of Mullins' work, can we build a network of teacher-scholars who will advance knowledge about the teaching/learning process? John Petras and James Hayes call attention to the fact that symbolic interactionists do not apply their perspective in the classroom: "interactionists do not appear to differ in any significant fashion from their colleagues as far as teaching techniques are concerned" (1973: 3). Extending the concern of Petras and Hayes, can different theoretical schools in sociology develop distinctive teaching/learning styles which thereby unite theory and practice? These questions are offered as openers. It is impossible to anticipate all of the issues which will confront the new breed of teacher-scholars who self-consciously scrutinize their craft.

Science is not always benign. Scholarly inquiry into the teaching/learning process may not yield quick pay-offs for the teacher. Indeed, the opposite may be the case. Scientific scrutiny of teaching and learning may be an invitation to stir up a storm of disquietude. Carefully researched questions will throw new doubts on cherished myths, thereby generating new anxieties. Two matters meriting further research illustrate some of the possible pitfalls awaiting the teacher-scholar.

The first concerns the oft repeated claim about the part sociological instruction plays in the cultivation of critical thinking. Ted Bradshaw and Sharon McPherron have recently conducted a national survey of sociology departments on curricular topics. According to their findings, the chief educational goal of the sociology curriculum is critical thinking. Few sociology teachers, however, have ever attempted to test empirically the attainment of this worthy goal. Charles Logan (1976) pioneered an empirical study which asked the obvious question: "do sociologists teach students to think more critically?" Logan's results are sobering, "These findings suggest that neither inclination nor ability to think critically or scientifically about statements on social issues is associated directly or substantially with increased levels of sociological training" (1976: 37). Logan comments on the implications of these unflattering results: "The professed concern among sociologists with teaching students to think more rationally and scientifically about social phenomena may be to a considerable degree lip service that masks a hidden curriculum" (1976: 41). Given the current mood of accountability in higher education, most teachers have not been eager to scrutinize their lofty goals. It seems safer to keep educational aims such as critical thinking in the murky swamp of untested idealism. In some cases, the present system of unverified claims serves well the partnership of inept teaching and cavalier complacency.

A second example of the disturbing effects of inquiry lies in the shift from an emphasis on teacher-centered activities toward more learner-centered activities. In many classroom settings, teacher-centered activities continue to enjoy popular support from both students and teachers. When teachers place greater emphasis on the students' academic work and the importance of achieving specified learning outcomes, the consequences may be lower popularity scores on teacher rating forms. I first came to this tentative conclusion a few years ago while evaluating my experimental work. Over a period of three semesters, I moved progressively away from the teacher-centered performances of lecturing and joke telling; the new instructional strategy placed more attention on specified learning tasks which were explored in small group settings. My popularity as an instructor suffered during this period. On scale items of "highly favorable" and "favorable," my personality slipped thirteen points, my knowledge of the field fell eight points, and my overall rating declined thirteen points. Such results forced me to reflect on the value of instructional innovation: "It was paradoxical that the instructor's intensive effort to create an alternative to mass instruction led to declining student respect of this learning endeavor" (1971: 316). Conversations with many faculty members who have experimented with learner-centered teaching models suggest similar experiences. In this age of student ratings administered by deans to help determine salary raises and tenure, it is not likely that faculty will jump for joy at the opportunity to place more responsibility on the student as learner. Instructional innovation in higher education appears to have reached a peculiar impasse in which many teachers can no longer afford risk-taking activities designed to maximize learning.

The Sociological Imagination and the Classroom

I am suggesting that scientific scrutiny of critical issues of teaching and learning will not lead to simple success or professional popularity. Yet, to be credible, teachers of sociology must make the effort. The integrity of teaching sociology requires practicing the central tenets of scientific inquiry. This means exercising the use of our most natural and immediate laboratory--through the experiences of our students--the classroom.

Sociology teachers often call upon C. Wright Mills and his vision of the sociological imagination. We never tire of quoting his challenge to transform private troubles into public issues. It is time to reassess Mills' challenge for our professional work. The private troubles of our classrooms must be transformed into the public issues of higher learning in a democratic society. We need to sustain the tension of a dual vision which looks inward to the learning struggles of students and outward to the requirements of informed public opinion. The tribal world of the classroom can also be recognized as a learning arena of civil discourse. Teachers of sociology who reflect critically on their tasks may find a new sense of intellectual excitement which reaches beyond the boundaries of the classroom.

This must have been the feeling of those sociologists who gathered in April, 1978, to exchange ideas--what we knew, thought we knew, needed to know if we were to improve our work as professors of sociology. Despite difficulties anticipated (and realized) these people explored a wide range of theory and research bearing on effective instruction. The chapters prepared for this conference appear on the following pages. They tackle some of the many questions about the ends and means of sociology instruction, and about the evaluation of the means to determine how well our ends are met. Thus we hope to have begun this task: to channel new energy into research which can provide a more substantial base of knowledge, so enabling us to do better the job which is so central for so many sociologists.

References

BAKER, P. (1971) "Alternatives to mass instruction in sociology." American Sociologist 6: 311-317.

BECKER, H. P. and H. BARNES (1938) Social Thought from Lore to Science. Washington, D.C.: Herren Press.

BRADSHAW, T. and S. MCPHERRON (1978) Data Report for Issues and Resources in the Undergraduate Sociology Curriculum (ASA Projects on Teaching Undergraduate Sociology). Washington, D.C.: ASA Teaching Resources Center.

BROWN, G. I. (1971) Human Teaching for Human Learning: An Introduction to Confluent Education. New York: Viking Press.

LOGAN, C. (1976) "Do sociologists teach students to think more critically?" Teaching Sociology 4: 29-48.

MULLINS, N. (1973) Theories and Theory Groups in Contemporary American Sociology. New York: Harper & Row, Publishers.

PETRAS, J. and J. HAYES (1973) "How symbolic interactionists can practice what they preach: Introductory sociology and knowledge as emergence." Teaching Sociology 1: 3-12.

WOLF, E. (1977) "Courtrooms and classrooms." Educational Forum 40: 431-53.

ZNANIECKI, F. (1940) The Social Role of the Man of Knowledge. New York: Columbia University Press.

Assessing the Ends We Serve: We Know Where We're Going Don't We?

Michael S. Bassis
The University of Rhode Island

Once upon a time a Sea Horse gathered up his seven pieces of eight and cantered out to find his fortune. Before he had traveled very far he met an Eel, who said,

"Psst. Hey, bud. Where 'ya goin'?"

"I'm going out to find my fortune," replied the Sea Horse, proudly.

"You're in luck," said the Eel. "For four pieces of eight you can have this speedy flipper, and then you'll be able to get there a lot faster."

"Gee, that's swell," said the Sea Horse, and paid the money and put on the flipper and slithered off at twice the speed. Soon he came upon a Sponge, who said,

"Psst. Hey, bud. Where 'ya going'?"

"I'm going out to find my fortune," replied the Sea Horse.

"You're in luck," said the Sponge. "For a small fee I will let you have this jet-propelled scooter so that you will be able to travel a lot faster."

So the Sea Horse bought the scooter with his remaining money and went zooming through the sea five times as fast. Soon he came upon a Shark, who said,

"Psst. Hey, bud. Where 'ya going'?"

"I'm going out to find my fortune," replied the Sea Horse.

"You're in luck. If you'll take this short cut," said the Shark, pointing to his open mouth, "you'll save yourself a lot of time."

"Gee, thanks," said the Sea Horse, and zoomed off into the interior of the Shark, there to be devoured.

The moral of this fable is that if you're not sure where you're going, you're liable to end up someplace else, and not even know it.

- a fable[1] -

"Would you tell me, please, which way I ought to go from here?"
"That depends a good deal on where you want to get to."

- Alice in Wonderland -

Teaching is a purposeful, goal directed activity. Teachers hope and expect that their students will take something away with them from their classroom experiences. Moreover all of the thought and effort that teachers spend in designing and conducting courses is grounded in notions, sometimes explicit and sometimes implicit, about what it is that students should take away. In simple terms we teach in terms of goals, ends, or purposes.

What should college students take away with them from their encounters with sociology? The discipline is too broad and our students are too varied to tolerate any prescriptive set of goals. What might be important for one instructor and appropriate for a given set of students might be meaningless for others.

Nonetheless an absolutely fundamental component of course or curriculum planning is the explication of goals. If we are clear about what it is we wish to accomplish, if we have thought through what it is we want our students to learn, not only do we have some basis for making decisions about how to reach our instructional goals, but we are in a better position to assess the extent to which we have reached them. The central point is that teaching and learning sociology can become more rewarding and productive activities as instructors become more self-conscious about the ends they serve. While each of the following four chapters was written to identify, assess, and perhaps contribute to knowledge that might be of use to instructors and departments seeking to clarify their instructional goals, each is predicated on the assumption that attaining such clarity is crucial and logically prior to all other activities in the instructional process.

In the first of these chapters Robert Stauffer explores sociology's contribution to a liberal education. To what extent are the goals of a liberal education identical to or at least compatible with a variety of instructional goals for sociology? What can sociology contribute to the liberal education of undergraduates and how might it maximize its potential contributions? In the process of reviewing the literature on these questions Stauffer raises important issues about the content and structure of sociology courses and about the training of undergraduate instructors.

Recently, various social trends have led instructors and departments to ask a new set of questions. Beyond instructional and curricular goals appropriate for students seeking a liberal education (and beyond the goals appropriate for graduate level training in sociology) what other sorts of goals are indicated for the new breed of student seeking vocational skills? In their chapter Professors Perrucci, Potter, and Huff report about an undergraduate concentration in applied sociology that they recently have helped develop at Purdue University. Here, quite in contrast to the liberal education goals described by Stauffer, students are expected to learn about human service organizations and to develop skills in data gathering, analysis, and program evaluation. The overall curricular goal is to prepare students for positions in public and private organizations in the human service sector. Again the message is clear. Decisions about what to teach and how to teach it are easier to make once we have carefully thought through and explicated our instructional ends.

But no matter whether our goals lead to liberal, preprofessional, or vocational education, they are all crosscut along four dimensions. Charlotte Vaughan's chapter points out that cognitive, effective, experiential, and interpersonal learning represent very different types or domains of learning and that each is relevant to undergraduate instruction in sociology. In addition, each domain includes a hierarchy of levels of learning that vary in complexity. Vaughan reminds us that knowledge of the major taxonomies of domains and levels of learning can help us recognize the implications of teaching and learning at one or another level and in one or another domain.

Finally, a chapter by Charles Goldsmid presents a clear and cogent case for articulating one's instructional goals in statements that express, in measurable terms, some knowledge or skill the instructor wishes students to learn. Goldsmid reviews the recent research and argument over the use of instructional objectives and concludes that goals left vague and general are inadequate guides for action--inadequate for teachers seeking to select appropriate teaching methods, learning activities, and evaluation techniques; for students seeking to focus their attention and energy more productively; and for colleagues seeking to share expectations about what students should be learning. Specificity lends clarity and clarity is essential to the demystification of the teaching and learning process.

There are those who will argue that teaching is an art, or at the very least a craft, and that creativity and spontaneity, so essential to effective teaching, can only be diminished through the mechanistic specification of objectives. While they are certainly explicit and measurable, instructional objectives are so narrow that they tend to trivialize the educational process. At best, or so the argument goes, students will learn much about the trees yet will remain ignorant of the forest.

The point is well taken. A good many sociology instructors, especially those who teach within the perspective of liberal education, are more comfortable specifying their instructional aims or purposes with general statements of overriding intent. While they may be less explicit such statements are more elegant and can more easily tap the consequential and the integrative aspects of learning.

An intense debate rages among those who demand more or less specificity in their statements of instructional intent. Clearly there are advantages and disadvantages associated with each alternative. However a technique does exist for articulating what students ought to learn that gives equal weight to the forest and the trees; to the consequential and the measurable, to the integrative and the specific. Proponents on both sides of the current debate will be challenged to reexamine their positions by reading Charlotte Vaughan and Richard Peterson's description of their synthetic approach (1975).

Debates will continue. In fact, they should be encouraged. But no amount of debate ought to obscure the fact that a heightened self-consciousness about the ends we serve can only improve the teaching of sociology.

Note

1. Mager, R. (1962) Preparing Instructional Objectives. Lear Siegler, Inc./ Fearon Publishers. vii.

References

MAGER, R. (1962) Preparing Instructional Objectives. Lear Siegler, Inc./ Fearon Publishers. vii.

VAUGHAN, C. and R. PETERSON (1975) "Introductory sociology: a behavioral objectives approach." Teaching Sociology 3: 6-20.

Sociology and Liberal Education: Some Observations for the KAKN Work Group on Assessing the Ends We Serve*

Robert E. Stauffer

Kalamazoo College

The goal, surely, of what we term general or liberal studies in colleges and universities is the creation of educated men and women. And while consensus regarding the full meaning of "educated" is, at least today, perhaps impossible to achieve--see, for example, the various conceptions summarized in the Carnegie report on the *Missions of the College Curriculum* (Carnegie Foundation, 1977)--I think there would be wide support for the following description. Let us define the educated person as one who possesses the following traits:

1. Knowledge of the general meaning and significance of, and, as importantly, the interrelationships among, the "big ideas," historical and modern, which characterize the civilization, at least, of which the individual is a part.
2. A capability to ask relevant questions about contemporary political and moral issues. This presumes both familiarity with history (Are the issues new? Have similar issues existed in the past and how were they dealt with? How have the issues developed?) and other contemporary societies (How unique are the issues? If not unique, how have other societies responded to them?). It also implies some literacy in modern natural science.
3. An aesthetic sensibility--an ability to recognize and derive pleasure from creative works and the creative process. This applies not simply to the humanities and fine arts but to scientific and social scientific scholarship as well.
4. An ability and a propensity to analyze arguments and assertions logically or critically.
5. The competence to communicate, orally and in writing, clearly and with at least some grace.
6. The development of an informed self-consciousness, or, to use an overworked but still meaningful notion, of an examined life.

7. A desire to continue to learn. (This trait was of course Mills' goal when he wrote that "We must try to end with men or women who can and will by themselves continue what we have begun: the end product of a liberal education . . . is simply the self-cultivating man and woman" (Mills, 1963: 369-370).

These traits obviously constitute an ideal type. We all know that liberal education regularly and necessarily falls short of achieving these ends. (And, incidentally, when students do approximate this type, it often is due only in part to formal education.) Further, it is undoubtedly true that many would find this list incomplete or wrongly weighted. And of course setting forth such a description does not begin to deal with questions such as the vocational worth of liberal education or the very difficult issue of the relationship between liberal education and moral values or political practice. Rather, the only purpose in beginning with a broad description of the liberally educated person is to provide a common reference or context for the ensuing discussion.

In consideration of liberal education, the central problems emerge not on the level of abstract goals or intentions but rather when one attempts to link these goals to more specific content and structures or formats. Invariably, at this point, the multiplicity of academic vested interests comes into play--specialists defending the particular value of their specialities, departments defending the particular value of their disciplines, sometimes whole faculties defending departmentalism against interdisciplinary initiatives. But resistance to more concrete consideration of liberal education is not only political in this self-interested sense; of equal importance, I think, is the relative lack of serious attention given to the possible relationships between the content or substance of diverse bodies of knowledge and skills, on the one hand, and the objectives of liberal education on the other. Too little convincing intellectual groundwork has been done, in other words, especially in the sciences and social sciences, that might result in more genuine interest in liberal education and greater openness to intellectual and structural changes intended to further that end.

The purpose of this chapter is to assess the available literature on the contributions one discipline--sociology--can make to liberal education, and to suggest some of the issues on which further thought and research are needed. In considering these matters, incidentally, I will not distinguish between single sociology courses (taken as electives or to meet distributional requirements) and the major in sociology. Even if it were advisable (which I doubt along with Joseph Zelan /1975: 197-198/), for prospective graduate students to receive a miniature version of this training in an undergraduate sociology major, it is also true that this would apply to very few of our contemporary majors. Thus, it seems more reasonable to consider the entire undergraduate experience as one of liberal education, with the major justified not as specialized training but rather as one mode of lending an important dimension of depth to this education.

Sociology and Liberal Education

A review of two probable sources suggests indeed that relatively little attention has been given to sociology and liberal education. In perusing some ten catalogs from renowned liberal arts colleges, I was able to turn up virtually no statements in the sociology sections about the meaning of this discipline for liberal education. Further, the lengthy bibliography on "Resources" prepared by Charles Goldsmid for Teaching Sociology (Goldsmid, 1976) contains only six references under the section on "Sociology and General Education," making this one of the smallest sections in the bibliography. Certainly, much of the growing literature on teaching sociology bears on a liberal education, but there is little which confronts the issue directly. My impression, in short, is that with a few important exceptions, sociologists--even those particularly concerned with undergraduate teaching--have not asked, at least in a formal way, what we are really about in our ostensible roles of liberal educators.

But there are exceptions--essays or commentaries which both can provide a basis for further discussion and can clarify the points on which disagreement is probably inevitable. Some of these draw from more sustained arguments or offer only tangential, although useful, points (Berger, 1963; Cuzzort and King, 1976; Mills, 1963; Page, 1949; Smelser, 1973; Wilson, 1971), and I will draw on some of these when appropriate. Three sociologists have confronted the issue more thoroughly and directly, however, and it is to them whom I want to turn now.

Perhaps the best known essay in this regard is Robert Bierstedt's "Sociology and General Education" (1963: 307-308):

> . . . like history in particular, the study of sociology liberates the student from the provincialisms of time, place, and circumstance and frees him from the constrictions of his natal culture. Secondly, it introduces him to the role of logic and of scientific method in the acquisition of knowledge and thus contributes, thirdly, to his sense of order and to his methodological sophistication. In the fourth place, sociology is a discipline that spans two cultures, the scientific and the humanistic, using as it does the method of science to explore the concerns and affairs of humanity. In the next instance, the fifth, I suggest that sociology initiates and keeps at the front of student awareness the ancient problem of the relationship between society and the individual. As a sixth point I referred to another ancient problem, the meaning of history, and declared that the philosophy of history, when it becomes positivistic, is indistinguishable from sociology and that it is the responsibility of sociology ultimately to find an answer to this agelong quest. Finally, in opposition to those who accuse sociologists of stylistic inadequacies, I maintained that the literary quality of their work is not one whit inferior to that which can be found in other learned disciplines.

While each of these contributions takes on added clarity and richness in the text of Bierstedt's predictably graceful essay, most are relatively straightforward--if controversial, as we shall see. Points two and three, however, need some explication here. "The abstraction of inquiry," Bierstedt's second point in the text, concerns abstract reasoning (thinking in terms of, and logically relating, abstract concepts) and, as consequence, learning "the nature and character of definition, and the rules of measurement" (Bierstedt, 1963: 45). Point three, the "sense of order," has to do with learning to find patterns behind appearances--to arrive inductively at orderly structural images. On both points, Bierstedt argues that unlike the humanities (which are seldom concerned with specific methodology) or the natural sciences (which simply assume it), sociology is useful since it forces upon the student a conscious awareness of and involvement with basic methodological problems and experiences.

What is striking overall about Bierstedt's approach is the range of implications he finds in sociology for liberal education. Indeed, as the following comparison suggests, the diverse impact he ascribes to sociology is such that most of the traits of the educated person described at the beginning of this chapter can plausibly be viewed as enhanced through sociological education.

Bierstedt's Assertations about:	especially relate to	Traits of an Educated Person
Liberated Mind		Relevant Questions about Political and Moral Issues Informed Self-consciousness
Logic and Scientific Method		Analyze arguments logically and critically Aesthetic sensibility
Two Cultures		Knowledge of "big ideas," especially interrelationships
Person and Society		Relevant Questions about Political and Moral Issues Informed Self-consciousness
Sociology and History		Knowledge of "big ideas" Relevant Questions about Political and Moral Issues
Prose of Sociology		Aesthetic Sensibility Competence to Communicate

The second argument to be considered was formulated partially in explicit reaction to Bierstedt. Marvin Bressler, writing on "Sociology and Collegiate General Education" (1967) believes our major task is to discover what, if any, impact sociology has on students' intellectual lives. This means, first, that we should attend more to what is unique to sociology than on what sociology may share with other disciplines. Thus, in commenting on Bierstedt, Bressler (1967: 55) writes that

> The author makes only one claim for the distinctiveness of sociology, its capacity to unite the "two cultures." Elsewhere in the essay he refers to this as "almost uniquely" a sociological contribution. Most of the virtues that Bierstedt ascribes to sociology inhere in other disciplines and are sometimes more appropriately identified with them. Students who are denied the delights of sociological prose will suffer no irremediable insult to their aesthetic sensibilities. Similarly, if, "like history in particular," sociology is antagonistic to parochialism, this antidote would presumably survive even in the absence of sociological instruction.

Secondly, Bressler is worried that Bierstedt represents only one perspective in sociology and one largely lacking in American sociological instruction--the sociophilosophical quest for historical meaning. As Bressler notes, seemingly without dismay, "the overwhelming number of courses in (American) sociology proceed as if the United States exhausts space and the present is the beginning of time" (Bressler, 1967: 57).

Preferably, then, according to Bressler, we should ask what it is that sociologists offer that is both unique to sociology and common to all sociologists. This, he asserts, is a set of epistemological emphases (objectivity, skepticism, categorization, abstraction, probability, relativism, context, and paradox) which represent science in general but which take on greater salience when applied to human affairs. (While a full equation would be overly facile, Bressler seems to be referring here to what Bierstedt called logic and scientific method, sense of order, and the spanning of the two cultures.) However, Bressler argues forcefully that were we able to communicate this social scientific ethos, we would also affect deeply the student's self and self-perception. The result, he believes, would not be the liberated activist, as Mills, for example, imagined (1963); but rather a more ironic, less naïvely individualistic, and more cautious person. But, in any event, Bressler believes the current need is to discover whether such an epistemology is imparted and to study whether this has any larger effects. These goals, he argues, must be prior to a more normative discussion of sociology and liberal education.

Bressler shares with Bierstedt an inductive approach, insofar as each writer purports to describe what sociology courses possibly do contribute to liberal education. The difference is that Bressler is (a) skeptical of some of Bierstedt's perceptions and (b) wishes to test hypothesized effects (thus limiting his list not only to what he thinks is likely but also to what is unique). The reader, however, does gain an impression from the tone of each essay that the authors tend to approve of their respective observations and thus would probably support rather different curricular models.

The third, relatively systematic approach to sociology and liberal education is found in a series of papers by Donald Levine, particularly his "Facing Our Calling: Sociology and the Agon of Liberal Education" (1971; see also Levine, 1966, 1970, 1973). Asserting strongly that sociologists have not examined adequately their undergraduate educational mission, Levine proposes a very different strategy from the two previously discussed.

Levine begins by noting a common theme regarding the value of sociology in writings of Comte, Marx, Durkheim, Weber, Mannheim, and Peter Berger; viz. that sociology offers "a unique means of de-mystification and intellectual clarification" (Levine, 1971: 3). The problem, however, is to identify the particular intellectual processes by which this de-mystification and clarification--this transcending of common sense culture--can occur through sociology. What intellectual skills, in other words, promote these goals, and to what extent does or can sociology's inclusion in a curriculum promote such skills?

Rooting his argument in cultural evolutionary terms, Levine refers to a series of differentiations among (1) the kinds of meanings culture provides; (2) the objects of orientation; and, (3) the separation of cultural from social institutions. Resulting from this process in the past century was "the cultivation of cognitive orientations to social objects pursued for their cultural value" (Levine, 1971: 6), or in essence, social science. But these cognitive orientations are themselves disciplines which Levine provisionally identifies as empirical methodology, categorical theory, normative philosophy, scientific reportage, utopiology, explanation, criticism, diagnosis, and prescription. (Some of these are either combinations of others or else more complex sets of other subdisciplines.)

Each of these disciplines, Levine suggests, ought to be included in liberal education, and this approach thus affords a means for examining and establishing curricula. Commonly today, he notes, sociology (and I presume some of the other social sciences) includes empirical methodology, categorical theory, scientific reportage, and explanation. But the point of the argument is not to defend or to expand sociology's present domain. Rather, Levine means to suggest that we begin with the need to teach these "social disciplines" and then arrange courses within the social sciences to assure that all of these disciplines or skills are, somewhere, systematically taught.

Levine, then, while perceiving current sociological contributions similarly to Bressler (and, less inclusively, to Bierstedt), nevertheless takes a different tack. To oversimplify, Bierstedt states that sociology offers a great deal; Bressler thinks sociology offers something of value and wants to find out; and Levine says what more generally the social sciences ought to offer and asks how sociology does or can contribute to this.

Other commentators on this general issue of sociology and liberal education generally reflect one or another of the positions already described. Let us note here only two examples. Everett Wilson, for example, in his discussion of "some uses of sociology," follows Bierstedt rather closely (Wilson, 1971: 599-602). The study of sociology, Wilson suggests, contributes to the discovery of self and the variety of the social world; to one's methodological acuity; to increased understanding of a wide range of phenomena (from business to literature); and to liberation from the customary, from the immediate (in time and space), and from the artificiality of "intellectual fissures."

Cuzzort and King, in a humorous if somewhat abstruse chapter on "The Uses of Social Thought" (1976), take a more circumscribed stance. They suggest that sociology has little to offer to the construction of utopian images (other than a certain testiness about unanticipated consequences); that when sociology tries to serve administrative power, it ceases to be sociology; and that sociology proper necessarily has made little contribution to increasing human happiness. Rather, according to these authors, sociology's value "lies in the extent to which a knowledge of how human social orders operate can free the individual from some of the more repressive forms of thought characterizing human social institutions" (Cuzzort and King, 1976: 325). This goal, which they note sociology shares with the humanities, is of course similar to the idea of liberation noted by Bierstedt and Wilson, and it is also the theme Levine considers central in the sociological tradition. However, Cuzzort and King also argue that sociology (or at least a "humanistic" sociology) teaches a skepticism toward intellectual life as well, since it is in the nature of sociological thought that a scientific certitude remains an impossible goal. In this sense, the unique epistemological value of studying sociology noted both by Bierstedt and Bressler (and hinted at by Levine) is emphasized.

There does exist, then, a small but useful literature on which to build subsequent discussion and research concerning sociology and liberal education. And while there are important differences concerning the particular value of sociology in contributing to various ends, one can discern substantial agreement from the writers discussed that sociology contributes to (or ought to) a greater sense of the social nature of self; an awareness of the provisional nature of any particular pattern of social organization and thus a kind of liberation from common sense; and, above all, an enhanced capacity for critical reasoning, resulting especially from the unique way in which sociology bridges the "two cultures."

What needs to be done? Perhaps most obviously, we ought to pursue Bressler's recommendation and study empirically what does happen to students as a result of taking sociology. One recent investigation, however, gives very little reason for optimism. Seeking to discover whether varying degrees of exposure to sociology courses promoted greater critical thinking, Logan discovered that only a course carefully and explicitly focussed on this goal made any difference (Logan, 1976). But more such research is needed, employing diverse dependent variables related to the various ends of education we have discussed and, when possible, controlling for the influences of non-sociology course work.

More than evaluation research is needed, however. In part because findings such as Logan's are discouraging and in part because, as the foregoing suggests, only a few have even thought systematically about sociology's liberal education ends, much more attention to what we're doing and ought to do is called for. It is possible, indeed probable, that some of the goals we have noted might be better achieved were more effective teaching strategies and techniques employed. Yet, I think that of at least equal importance is the gap between the broad and general goals commonly espoused, on the one hand, and the content and structure of curricula, on the other. A perusal of college catalogs, for example, suggests that undergraduate sociology curricula essentially ape disciplinary subfields, raising the question whether much effort has gone into the liberal educational significance of the courses taught. Further, we know that textbooks are used almost universally in first courses of sociology, and very frequently in other courses (Bradshaw and McPherron (1978: 13). Yet many have questioned the text as an intellectual tool, especially at the introductory level, finding many texts to be mainly catalogs of disembodied concepts and contextless and easily-dated facts. Certainly, most introductory sociology texts convey the range of concerns in the discipline and thus something about the nature of the field. But this seems, at best, oriented more toward informing the student about possible future courses than toward providing liberal education _in_ the introductory course. Does not this again suggest the need to think more carefully about the relationship between our larger goals and the curricular materials we employ?

To be more specific, we need first to pursue more fully the kind of work reported on earlier in this chapter, particularly that of Bierstedt and Levine. That is, given or developing clear statements of the ends of liberal education, what can and should sociology offer? What within sociology's repertoire of ideas, skills, and works have the potential to contribute to specified traits of the educated person? For some provocative and germane ideas on the value of the sociological tradition, two important essays are Bierstedt's "Sociology and Humane Learning" (1974) and Robert Nisbet's _Sociology as an Art Form_ (1976). Bierstedt offers a rationale for what might be called a "great books of sociology" approach, and Nisbet emphasizes the value of treating the sociological "classics" as no less artistic or creative than parallel developments in the fine arts. But these are only examples, and the general point remains: What are our goals and how best can sociology serve those goals?

A second issue is very much related to the first. How ought sociology to fit into the larger general education curriculum? If, for example, Bierstedt were correct in his argument that a large number of liberal education goals _are_ served by sociology, along with other disciplines, does this mean that (a) sociology gains importance as a separate liberal discipline or (b) the various disciplines are unnecessarily duplicating each other and should be combined into interdisciplinary courses or programs? Conversely, if Bierstedt is wrong in thinking sociology _effectively_ achieves diverse goals (which is probable), then does this mean sociology ought to concentrate on possibly unique ends and improve its contribution in that regard (e.g., the epistemological emphases of Bressler or certain of the cognitive skills identified by Levine) or does it mean, again, that sociology's broader potential would be better realized if courses were offered in interdisciplinary settings?

An interesting model relevant to these questions has been suggested by Smelser (1973). Rejecting the "proliferating specialized courses that mirror the graduate curriculum" (1973: 149), Smelser suggests that the various social science disciplines do two things. First, each would develop two-year introductory courses, ideally team-taught, not to survey the discipline but to "expose the underlying assumptions of the discipline, its major theoretical frameworks, and the fundamentals of its stock of knowledge" (1973: 149). Following this would be not more specialized courses but rather courses dealing with "general ideas and problems," often interdisciplinary and team-taught. He mentions, among other examples, "The Implications of Bureaucracy for Human Freedom" and "Conformity and Deviance in Organized Social Life" (1973: 150). This scheme combines a sustained emphasis on critical reasoning and epistemological skills of the discipline with the possible benefits of using sociology in an interdisciplinary manner to serve the larger curriculum through courses more germane to students' interests, needs, and past and future experiences. The point of noting this example is, of course, not to suggest that it represents the ideal solution. It is only one possibility--and one with obvious though probably not insoluble problems. Rather, it merely suggests that when one does begin to ask how sociology may serve liberal education goals, interesting designs--and ones quite different from customary approaches--are possible and perhaps necessary. It is worth remarking that the same spirit which seems to inform Smelser's suggestions is found, at least in part, in the recent report on the Core Curriculum at Harvard (Rosovsky, 1978). With regard to social science courses, this report rejects "introductory courses that are a potpourri of what a department has to offer" and encourages courses which "give students a systematic introduction to some major social science theories or approaches and to the empirical foundations of those theories and approaches" (Rosovsky, 1978: 23). Furthermore, a concrete example of carefully linking course content to broader liberal educational goals is provided in Vaughan and Peterson's description of their approach to introductory sociology at Cornell College (Vaughan and Peterson, 1975).

The third and last issue I wish to raise about sociology curricula and liberal education is the education of the teacher. It is commonly noted that most future sociology teachers receive mainly a research education, although there has been some growth in the number of graduate courses on the teaching of sociology. It is less frequently observed, however, that the kinds of substantive knowledge needed by the teacher of undergraduate sociology may be rather different from that acquired either in graduate school or even on the job. If one begins not by saying that I will teach what I know but rather by asking what ought to be taught, the potential or actual limitations of our knowledge are highlighted. In particular, if a central purpose of teaching sociology in a liberal arts context is to demonstrate the value of the discipline for illuminating diverse realities and especially those about which students are learning in other courses (e.g., philosophy, psychology, history, literature), then is it not essential that the teacher be conversant with and indeed involved in intellectual concerns transcending the discipline? A similar point was made two decades ago by Page (1959: 589) when he wrote, in reference to teachers capitalizing on sociology's intellectual marginality, that

> Such teaching superiority requires the curiosity--and the time, alas--to explore sociology's border regions: to study Freud, for example, as well as Weber; to read, say Fortune and Dissent as well as the American Sociological Review; more, to know some philosophy and fiction and poetry as well as Parsons. Today, this plea for broadly educated sociologists may seem utopian. But surely teachers who are confined by their specializations and limited in their ability to reveal the diverse applicability of sociological analysis--who, in other words, do not exploit the advantages of marginality--restrict the educational possibilities of their discipline.

Needless to say, this problem becomes even more acute when sociologists participate in interdisciplinary courses and programs, for these can only succeed when materials are genuinely integrated, not when a series of specialists give "perspectives" on the issue.

The possible implications of responding to this problem in particular are several and difficult. Should the intellectual skills looked for in prospective teachers of undergraduate sociology be broadened or redefined? Should there be separate graduate programs for teachers? Would it, to go even further, be desirable to separate the undergraduate teaching and research enterprises institutionally, as Smelser suggests (Smelser, 1973: 153), such that undergraduate teachers would have the time and motivation to continue to educate themselves broadly? Surely no easy solutions are forthcoming. But also surely this remains a central concern for any serious consideration of sociology and liberal education.

I have attempted in the foregoing to indicate the major existing perspectives on the significance of sociology for liberal education and to lay out some of the seminal issues for further discussion, research, and experimentation in this regard. To conclude, I want to offer a more personal observation. An implication of this chapter is that we ought to translate, much more explicitly and systematically than we do now, the general ends of education into more specific curricular and course goals. Yet just as there are problems with unreflective teaching of the discipline, there are also potential hazards in determining and teaching to specific objectives. In both cases, the larger point of sociology as a liberal art can be obscured--that is the intellectual excitement nourished by the interplay of some very good ideas and the big issues of the real world. The loss of this excitement is all too obvious in many conventional, discipline-focused courses. Less obvious but still real is that an emphasis on teaching specified objectives faces the danger of distorting, both for the teacher and the student, this larger meaning of the sociology--the fact that its major ideas arose out of, and remain centrally relevant to, political and moral controversy. Granted, this enormous potential of sociology--to illumine the nature of public affairs and to make more thoughtful one's reaction to them--can itself only be conveyed successfully through careful curricular and course planning. But, and this is my point, communication of this larger meaning also involves, on the part of the teacher, and in the relationship between teacher and student, a probably irreducible chemistry of problem, ideas, and intellectual passion. This means at the least, that sociologists who concern themselves with undergraduate teaching not replace a narrow disciplinary expertise with a narrow pedagogical expertise. Rather, what is most needed, I believe, are sociology teachers who are intellectuals in the best and classic sense of the term--who live through ideas of political and cultural import and who communicate with conviction, rigor, and sensitivity the significance of this enterprise.

Note

*I am grateful for the helpful comments on earlier drafts of this chapter from Charles A. Goldsmid, Michael S. Bassis, and Charlotte A. Vaughan. And for their being regularly available role models of inspired teaching of sociology and political science respectively in a liberal arts context, I owe a real debt to my colleagues Kim Cummings and Franklin Presler.

References

BERGER, P. (1963) Invitation to Sociology: A Humanistic Perspective. New York: Doubleday-Anchor Books.

BIERSTEDT, R. (1963) "Sociology and general education," pp. 40-55 in C. Page (ed.) Sociology and Contemporary Education. New York: Random House. Also printed in Bierstedt, R. [ed.] (1974) Power and Progress: Essays on Sociological Theory. New York: McGraw-Hill Book Company.

--- (1974) "Sociology and human learning," pp. 309-321 in R. Bierstedt (ed.) Power and Progress: Essays on Sociological Theory. New York: McGraw-Hill Book Company.

BRADSHAW, T. and S. MCPHERRON (1978) "Issues and resources in undergraduate sociology curriculum: report to respondents." ASA Project on Teaching Undergraduate Sociology.

BRESSLER, M. (1967) "Sociology and collegiate general education," pp. 45-62 in P. Lazarsfeld et al. The Uses of Sociology. New York: Basic Books, Inc.

CARNEGIE FOUNDATION FOR THE ADVANCEMENT OF TEACHING (1977) "The mission of undergraduate education," pp. 150-163 in Missions of the College Curriculum. San Francisco: Jossey-Bass Publishers.

CUZZORT, R. and E. King (1976) Humanity and Modern Social Thought, 2nd edition. Hinsdale, Illinois: The Dryden Press.

GOLDSMID, C. (1976) "Resources." Teaching Sociology 3 (April): 339-378.

LEVINE, D. (1966) "The social sciences and liberal education." Lecture presented at the University of Chicago, October 6.

--- (1970) "Sociology confronts student protest." School Review 78 (August): 529-541.

--- (1971) "Facing our calling: sociology and the agon of liberal education." Denver, Colorado: Presented at the annual meetings of the American Sociological Association.

--- (1973) "The politicization of sociology: implications for liberal education." New York: Presented at the annual meetings of the American Sociological Association.

LOGAN, C. (1976) "Do sociologists teach students to think more critically?" Teaching Sociology 4: 29-48.

MILLS, C. W. (1963) "Mass society and liberal education," in I. Horowitz (ed.) Power, Politics, and People: The Collected Essays of C. Wright Mills. New York: Oxford University Press.

NISBET, R. (1976) Sociology as an Art Form. New York: Oxford University Press.

PAGE, C. (1959) "Sociology as a teaching enterprise," pp. 579-599 in R. K. Merton et al. (eds.) Sociology Today: Problems and Prospects. New York: Basic Books, Inc.

ROSOVSKY, H. (1978) "Report on the core curriculum," Faculty of Arts and Sciences, Harvard University.

SMELSER, N. (1973) "The social sciences," in C. Kaysen (ed.) Content and Context: Essays on College Education. New York: McGraw-Hill Book Co.

VAUGHAN, C. and R. PETERSON (1975) "Introductory sociology: a behavioral objectives approach." Teaching Sociology 3: 6-20.

WILSON, E. (1971) Sociology: Rules, Roles, and Relationships, rev. ed. Homewood, Illinois: The Dorsey Press.

ZELAN, J. (1975) "Undergraduates in sociology," pp. 183-198 in M. Trow (ed.) Teachers and Students: Aspects of American Higher Education. New York: McGraw-Hill Book Company.

Applied Sociology and Community Organizations: New Goals for Learning and Working

R. Perrucci, H. R. Potter, C. R. Huff
Purdue University

Historical Roots of Applied Sociology

The current interest in applied sociology which is evident in our field (Morrissey and Steadman, 1977; ASA Footnotes, 1978; ASA Workshop, 1978) and which recently has been manifested in proposals to revise sociology curricula at many colleges and universities is, in many ways, a reflection of some of the earliest interest of American sociologists. Contemporary debates concerning the respective merits of basic research vs. applied research, quantitative vs. qualitative methods, and traditional vs. more applied sociology curricula are certainly not new debates. What may be new about these discussions is that they are occurring in an era of rapidly declining enrollments, especially among undergraduate sociology majors.

It may be useful, then, to put in historical context some of the current issues surrounding undergraduate education in sociology. The historical development of applied sociology, and some of the tensions which have been associated with that development, have been documented by Lazarsfeld (1975). In tracing that development, three phases, or eras, seem to be of particular significance: (1) the social improvement phase, (2) the search for scientific status, and (3) the attempt to integrate sociological knowledge and applied work. Drawing on Lazarsfeld's (1975) work, as well as a more general reading of historical works in applied sociology, we shall briefly summarize this development before addressing more specific concerns.

Early American sociology was, of course, closely connected with a quest for social improvement. Many of the early sociologists were ministers, and the field was heavily focused on "social problems," such as crime, poor housing, poverty, and the assimilation problems associated with massive waves of immigrants. Coser (1956: 17) notes that the first textbook of American sociology (Small and Vincent, 1894) contained the statement, "Sociology was born of the modern ardor to improve society." This social problem/reform emphasis was even reflected in the first graduate sociology program in the United States (Columbia, 1894) whose catalogue stated, in part:

> It is becoming more and more apparent that industrial and social progress is bringing the modern community face to face with social questions of the greatest magnitude, the solution of which will demand the best scientific study and the most honest practical endeavor . . . It is in the city that the problems of poverty, of meandering, of intemperance, of unsanitary surroundings, and of debasing social influences are met in their most acute form. Hence, the city is the natural laboratory of social science (Lazarsfeld, 1975: 1-2).

These interests led to the early social survey movement, which involved the systematic collection of data concerning many of these social problems. By 1912, the Russell Sage Foundation had established a department of surveys and information; by 1928, that department had catalogued at least 2,000 social surveys.

By the end of World War I, there were a significant number of persons who had managed to acquire training in sociology. Many of them had been involved in the social survey movement, while others had obtained formal graduate training in the earliest American departments of sociology or in foreign universities. The post-World War I period was marked by a concern with the scientific status of sociology and an increasing emphasis on the development of rigorous research methods and statistical techniques. The development of the Social Science Research Council was central to this focus on methods, as exemplified in the intense debates over the scholarly merits of Thomas and Znaniecki's The Polish Peasant. Those debates continue today in the form of discussions of the relative merits of quantitative vs. qualitative methodology. However, the implications of this second phase are much broader than that debate. The general shift of sociologists in the direction of basic theoretical and empirical work was a significant contrast to their earlier action and reformist orientation. In fact, until about World War I, sociology and social work were quite similar in their interests, with both fields focusing on urban problems associated with immigration, urbanization, and industrialization. However, just as sociology turned toward more rigorous research methods in its search for respectability, social work sought professional prestige through the perfection of Freudian casework practice (Meyer et al., 1957). The two fields went in very different directions, then, and communications between them continues to disclose these historical tensions.

During World War II, the government increasingly used social research in tackling such problems as civilian morale, the training of soldiers, the strategies of enemy forces, and propaganda. Sociologists had also been extensively involved in the Department of Agriculture, especially in the development of land grant colleges, whose mission it was to improve the lives of American farmers. These activities established a linkage between sociology and government organizations, a linkage which is still apparent to some extent in federal agencies and in Presidential commissions. These developments, along with the work of Lewin and Moreno in developing action research on small groups, fostered increasing interest in the integration of social science knowledge and practical problems, even social structural problems. That interest is different from the interests of the early reformers, in that it emphasizes the application of empirical methods in analyzing field problems and it lacks the moralistic, missionary spirit of the earliest sociologists in the United States. Nevertheless, many of the current substantive concerns are quite similar to those expressed during that early phase.

Undergraduate Education in Applied Sociology

Undergraduate education in sociology traditionally has been concerned with one of two general objectives: to expose students from diverse fields to sociology as part of a liberal education (see the preceding chapter by Stauffer); or to expose students to sociology in order to recruit majors and to prepare them for graduate training. The somewhat incompatible goals of liberal and professional education have been joined recently by vocational education, which has all the popularity of the uninvited weekend guest. The traditional goals may well continue to meet the needs of some students, although currently it appears that their numbers will be smaller.

There are a number of reasons for an increasing interest in a vocational emphasis. First is the changing interest of undergraduate students themselves. Many of them are expressing greater interest in specific marketable career skills and opportunities represented in the bachelor's degree. Second is an early stage of concern for organizational survival on the part of departments, and universities in some cases, as declining enrollments threaten to reduce the number of faculty, courses, and programs. Both of these are pragmatic considerations. Intellectually, however, a number of developments within the discipline indicate an increasing interest in applied issues. Policy research has increased in many areas, and social indicators are being developed to expand our knowledge of changes in society. Although many of these efforts have been at the national level, they have diffused to the state and local level as well. At the national level, particularly where new methods are being developed, education through the advanced graduate level is often an appropriate prerequisite for work in these areas. Through past experience, the discipline has developed skills which can be transmitted to undergraduate students and which are very suitable for the kinds of program monitoring carried out by many local agencies.

The somewhat ambivalent feelings that seem to be associated with the movement of sociology departments toward vocational curricula are, in our view, based on a misunderstanding of the very practical aspects of sociological knowledge. An applied sociology is not simply a vocational effort that narrows, restricts, and reduces sociological knowledge to technical know-how. It is an intellectual learning experience intended to blend theory and practice, and in so doing deals with issues that are at the center of theoretical and empirical endeavors in sociology.

These issues involve both social structural and social psychological variables. The task of selecting appropriate variables for understanding patterns of societal, organizational, and individual behavior is often complex. This may be as true in the applied setting as in academia, unless it is assumed that only simplistic indicators are used in applied activities. Discipline-based theoretical and methodological problems generally contribute to this but, in addition, the specific applied setting imposes various restrictions on the usefulness of variables. Local or state personnel may be able to do little to modify the current birth rate or the increasing number of older persons in the society, but these national demographic trends may be very relevant to their program planning. The identification of a set of salient variables that describe clients or organizational performance, on which data are collected systematically, may be useful for monitoring an organization's activities; in addition, over time, these data may become valuable for studying trends. We see these conceptual and methodological issues as being vital to three distinct applied activities: (1) developing policies and programs, (2) developing data bases for monitoring programs, and (3) developing training programs.

In the last three years, the Department of Sociology and Anthropology at Purdue University has developed a new undergraduate concentration in applied sociology. The concentration is designed to have three distinct, interrelated components. Only the first component, the curriculum, is presently in place. The other two are in early stages of development and will be described as they are currently planned. The curriculum has 39 semester hours of course work, 21 of which consist of required courses in theory, applied sociology, social psychology, social organization, statistics, field methods, and evaluation research. The student then selects two options for the remaining 18 hours. These options reflect the particular setting in which the student would like to work (e.g., criminal justice settings, health care delivery organizations, social welfare agencies) and are designed to provide a firm basis of knowledge about the nature of those settings and the interpersonal and organizational problems involved in serving clients.

The concentration in applied sociology grew out of two considerations: first, as described above, is the growing need for undergraduate curricula for students interested in the application of sociology to public policy and human problems; second, the increased demand for public accountability in public and private organizations, which requires them to monitor and evaluate their programs and to collect and evaluate data essential for eligibility for federal programs and compliance with federal regulations.

Undergraduate students in applied sociology will be prepared with knowledge of human service organizations, and with special skills in data gathering, analyses, and program evaluation. It is expected that students will be prepared for positions in public and private organizations concerned with public accountability. The strength of the program lies in its emphasis on organization-centered skills that prepare a graduate for such activities as program evaluation, client surveys, organization-community relations, and inter-agency relations essential to organizational programs.

The applied nature of this concentration is somewhat similar to social work education, which prepares its graduates for positions in casework, group work, community organization, administration, and policy formulation. However, the applied curriculum incorporates a greater emphasis on research skills and a more thorough grounding in selected substantive areas of sociology.

The second component of the applied sociology concentration is a summer internship experience following the sixth semester in the program. The internship is based upon and follows three courses: one course in applied sociology and two companion courses that are team taught and taken during the sixth semester.

One course (Case Studies in Organizational Analysis) is designed to give students exposure to the everyday operations, short term and long term goals, and special needs and problems of a variety of human service organizations. This will be accomplished mainly through the use of case studies of particular organizations and agencies, and through the active participation of agency personnel in the course. For example, students may read materials concerning the development and operation of community mental health facilities. Such materials would include an examination of legislation and Federal regulations, problems of interorganizational coordination, approaches to assessing the effectiveness of such Centers, staffing problems, and techniques for developing links between Centers, supportive community facilities, and families of patients. In addition, we expect to have a senior staff person from a community mental health facility participate in the course to give his or her point of view on the short term and long term needs and problems of such Centers.

The Case Studies course will include materials on community planning agencies, mental hospitals, and prisons or juvenile justice agencies. Students will be exposed to different perspectives on life in such organizations through the use of written case materials by social scientists, films such as the Wiseman documentaries, and books or interviews by ex-inmates, patients, welfare recipients. Finally, as indicated above, agency personnel will provide students with their perspectives on problems and needs of their organizations.

The second course (Field Studies and Evaluation Research) has a number of related objectives. It will provide students with an introduction to program evaluation approaches and techniques. It will also allow students to design an evaluation study based on a realistic problem put forward by agency participants in the Case Studies course. Finally, it provides a preparation for the summer internship, in that students will attempt to conduct small scale evaluation studies during this time. The level of the course is aimed at the type of in-house evaluation that local and state agencies undertake. It will also give students experience in identifying program goals, analyzing data, and writing reports on a small scale. Topics to be covered include (a) identifying program goals, stated objectives, and services provided; clientele served and potential clientele; specified indicators of outcomes; and variables affecting outcomes; (b) types of evaluation designs: descriptive surveys, ex post facto descriptions and other non-experimental designs, experimental and quasi-experimental designs; (c) validity of indicators; (d) sources of data: existing records and reports, interviews, questionnaires, and observation; (e) sampling; (f) analysis of data with emphasis on clarity and simplicity; (g) reports, both written and oral; and (h) problems: changing programs and goals, the evaluator's role, gaps between data based recommendations and policy decisions.

The third course (Applied Sociology) is an integral part of the Concentration in Applied Sociology and is designed to provide students with an opportunity to examine the uses of sociology in fields such as law and criminal justice, medicine and health services, social welfare and social work, mental health, and community organization and planning. The first section of the course focuses on the historical development of applied sociology, analyses of the various roles which an applied sociologist might play, and some of the major issues confronting sociologists working in applied areas of research and practice. The second section of the course consists of analyses of various case illustrations in which sociological concepts and perspectives have been applied, successfully or unsuccessfully, in areas such as those mentioned above.

During June and July of the summer session following the sixth semester students will participate in an eight week internship in an agency or organization in the region. Each student will have prior familiarity with the settings in which he or she is placed. This familiarity will stem from two courses taken during the sixth semester and contact with agency personnel who participate in the Case Studies course, and by the faculty involved in the three courses described above.

During the seventh semester, following the internship, there will be a post-internship seminar. The objectives of the seminar are to (a) evaluate the utility of the courses in the Applied Sociology program for the internship experience; and (b) examine the results of the student evaluation projects.

It should be clear that the teaching goals emphasized in this program are multiple in nature. Following Bloom et al. (1956), Vaughan and Peterson (1975), and Vaughan (1978), we are interested in moving beyond the transmission of concepts, theories, and methodologies (the cognitive domain), to achieve higher order goals associated with application of knowledge in a problem-solving context. For example, several courses in the curriculum (Case Studies in Organizational Analysis, and Field Studies and Evaluation Research), and the summer internship require students to examine and compare their knowledge of agency structures and problems with that provided by agency personnel. This will require students to draw upon a wider range of critical skills and develop broader understanding of organizational phenomena than is usually expected in courses that stress cognitive objectives. Both the utility and the limitations of sociological theory and method should be revealed by the confrontation of theory and practice.

The final component of the program is a job placement program designed to assist graduates of the applied sociology concentration in finding positions related to their education. Job placement will require us to establish contacts with agencies and organizations in our region, informing them about our program and the ways in which graduates can be valuable additions to their organizations. In some cases this will be difficult, as many human service organizations are not accustomed to thinking of sociology graduates as potential employees, or have not yet recognized the potential value of creating new positions. However, human service organizations increasingly are faced with the need to (a) monitor and evaluate their programs; (b) anticipate their future needs and program changes; (c) anticipate changes in size and composition of their clientele; and (d) collect and evaluate a variety of data needed to comply with changing governmental regulations, and to be eligible to participate in new and continuing federal programs.

As a first step in the job placement effort we have carried out two surveys designed to assess the marketability of a baccalaureate degree in sociology; one concentrated on the potential employer's perceived needs, and the other on the graduate's needs on the job (Irving et al., 1978). The employer survey included 400 social service and government agencies from the ten largest cities in Indiana. In addition, in a national sample, 100 business and industrial corporations were selected from the 1978 College Placement Annual, from a list of companies that have hired social science and liberal arts graduates in the past.

Employers were asked how many persons their company or agency employed (including both those with and those without degrees in sociology), the positions for which they would hire sociology graduates, and whether they project an expansion in the job market for graduates with a Bachelor's, Master's or Doctorate in sociology. They were also asked to rate various subjects (e.g., statistics, theory) and activities such as interviewing or counselling, as to their importance for work on the job, and to rate the relative preparedness in these areas of past sociology graduates they had hired.

For the alumni sample, questionnaires were sent to 170 graduates having received their degree in sociology from Purdue within the last five years. The alumni were asked basic demographic questions about their age, sex, and residence, whether they were working (full or part-time), attending college or in the military, and then were asked specific questions about where they were employed, what their activities were on the job, and what their plans were for future schooling. They were asked to rate the same subjects and activities as did the employers, except that employees were asked about extent of use of these skills on the job (within the past week, past month, entire career or never). For the same subjects and activities they were also asked to rate their past training and whether they recommend that other sociology graduates should have some background in the specified areas.

The subject areas included in both surveys (and typically parts of traditional sociology curricula) include: statistics, computer programming, writing skills, interpersonal communication skills, research methodology, program evaluation skills, economics, and political science, community planning, knowledge of other cultures, sociological theory, criminology and juvenile delinquency, human development and psychology, social work practice techniques, social work policy, field experience/on the job training, and ability to communicate in more than one language. The activities included in both surveys were: interviewing, writing proposals or reports, literature research, public relations, supervision, counselling, typing and filing, and teaching or training.

This preliminary survey provides us with some idea of the particular subjects and activities that employers and graduates believe to be important for employment. In general, the results from both the employer and alumni surveys indicate that the skill and subject areas most frequently cited as important on the job are writing, interpersonal communication skills, public relations, human development and psychology, and the ability to write proposals and reports.

Subjects in the sociology curriculum that most alumni (75%) recommend should be included are statistics, writing skills, knowledge of other cultures, and interpersonal communication skills. At least 50% of alumni also recommend sociological theory, criminology and juvenile delinquency, social work practice techniques and policy, interviewing, writing proposals or reports, and field experience/on the job training.

The general conclusion drawn from this survey is that what we must do "is to better integrate sociological theory and methods with a wider variety of occupational positions. Apparently students feel that they are graduating and entering the world of employment (or unemployment) without knowing what to expect or for what types of jobs they are qualified" (Irving et al., 1978: 14).

Conclusion

Sociology departments whose faculties wish to consider the development of programs in applied sociology will be confronted with a number of important issues. In recognition of this fact, the American Sociological Association and the Pacific Sociological Association jointly sponsored a workshop on teaching applied sociology. This workshop, which was held on April 11-12, 1978, in Spokane, Washington, focused on several of these critical issues, including:

(1) concepts and definitions of applied sociology;
(2) a national survey of internships in applied sociology;
(3) curriculum resources;
(4) field placement and evaluation of students and agencies;
(5) teaching applied research skills; and
(6) applied sociology from the agencies' perspective.

Workshop participants from diverse colleges and universities, as well as human service agencies, shared their ideas about, and experiences with, the development and operation of applied sociology programs.[1]

We believe that an applied sociology program is compatible with many undergraduate programs in sociology, including the existing curricula. The program described here was developed in large part by using the existing curriculum. Courses in juvenile delinquency, criminology, sociology of medicine, complex organizations, and community, as examples, are often organized so that their content is very appropriate for students with applied interests. Obviously, not all departments are inclined to include such a program in their undergraduate curriculum, and we recognize that the test of our program still lies ahead, as its graduates seek jobs.

Note

1. While the entire workshop cannot adequately be summarized in this chapter, readers interested in obtaining further information may wish to contact the workshop coordinator, Professor Joseph DeMartini, Washington State University, Pullman.

References

AMERICAN SOCIOLOGICAL ASSOCIATION (1978) Workshop on Teaching Applied Sociology. Spokane, Washington. April 11-12.

AMERICAN SOCIOLOGICAL ASSOCIATION (1978) Footnotes: 6: 1-7.

BLOOM, B. et al. (1956) Taxonomy of Educational Objectives, Handbook 1, Cognitive Domain. New York: David McKay Co.

COSER, L. (1956) The Functions of Social Conflict. New York: Free Press.

IRVING, L., HUFF, C. and J. STAHURA (1978) Putting Sociology to Work: An Assessment of Graduates' and Employers' Needs for Improving the Occupational Relevance of Sociology. Department of Sociology and Anthropology, Purdue University.

MEYER, H., LITWAK, E., THOMAS, E. and R. VINTER (1967) "Social work and social welfare," in P. Lazarsfeld, W. H. Sewell, and H. Wilensky (eds.) The Uses of Sociology. New York: Basic Books.

VAUGHAN, C. and R. PETERSON (1975) "Introductory sociology: a behavioral objectives approach." Teaching Sociology 3 (October): 6-20.

VAUGHAN, C. (1978) "Identifying course goals: domains and levels of learning." (unpublished)

Identifying Course Goals: Domains and Levels of Learning*

Charlotte A. Vaughan

Cornell College

An important component in successful teaching in sociology or in any other discipline is identifying what it is that we wish to accomplish through the teaching process. All too often we may operate in terms of implicit goals or hazy objectives, without giving the matter much thought. In fact, it may be no easy task to decide precisely what it is we want to accomplish through our various sociology courses. Do all our main goals lie in teaching content? What content, and how much? To what purpose? Do we want to teach thinking skills? Which ones, and to what extent? Do we want to teach such things as appreciation for the sociological perspective? What is "appreciation"? Are there other kinds of learning that we want to promulgate?

Taken seriously, these can be perplexing questions. There are perhaps as many answers as there are sociology instructors. I do not intend to supply an answer but rather to provide information about some important resources that can help us in the process of developing our own answers. These resources include attempts to identify different types of learning, such as intellectual or attitudinal, and the various degrees that can be attained within each type. They describe the levels of intellectual learning from simple to complex and identify the various stages in internalizing attitudes. These taxonomies display the range of learning alternatives from which we can then make informed decisions about our own desired teaching outcomes.

Making conscious decisions about course goals puts us in greater control of the teaching process, since we can then plan strategies specifically designed to meet them. Translating hazy goals into recognizable objectives gives direction both to choices of teaching strategies and to evaluations of students' progress--so assessing course efficacy and instructor's success. The case for translating course goals into instructional objectives has been made elsewhere.[1] This chapter is concerned with the problem of selecting goals themselves, although it will lead toward developing useful instructional objectives.

Domains of Learning

In 1948 a group of College and University Examiners embarked upon the process of developing a classification of educational objectives. Defining educational objectives as "explicit formulations of the ways in which students are expected to be changed by the educative process" (Bloom et al., 1956: 26), they referred to the major *types* of objectives as "domains." They identified three domains--the cognitive, affective, and psychomotor.

The *cognitive* domain includes educational objectives concerned with knowledge and intellectual abilities. The *affective* domain includes objectives "which emphasize a feeling tone, an emotion, or a degree of acceptance or rejection . . . [They are] expressed as interests, attitudes, appreciations, values, and emotional sets or biases" (Krathwohl et al., 1964: 7). *Psychomotor* objectives involve motor skills and manipulation of objects.

While major efforts at building typologies of educational objectives have been in these three domains, some other domains have been identified. They include the perceptual, experiential, and interpersonal. The *perceptual* domain has been defined as sensing, attending, distinguishing, and recognizing (Tuckman, 1972).[2] The *experiential* domain crosscuts the cognitive, affective, and psychomotor, treating exposure to and participation in experiences as including all three. *Interpersonal* learnings also crosscut other domains, stressing changes through relationships with others in a group setting.

Of these six domains, the cognitive, affective, experiential, and interpersonal appear relevant to teaching sociology. While we sometimes engage students in the psychomotor tasks of operating a calculator or computer, typically we are not concerned with developing psychomotor learnings as such.[3] Similarly, we tend to assume perception or to develop it through the cognitive domain. Therefore, we will consider these two domains further only as they impinge upon the other four.[4]

The Cognitive Domain

The domain of most concern to us as educators in sociology is the cognitive domain. A perusal of goal statements in published sources, summarized in Classification 1, shows that the major stated goal is that of transmitting to students a body of knowledge, including the nature of the perspective, concepts, theories, data, and research methods. The other goals are also heavily cognitive: developing understandings and insights, contributing to general education and thinking abilities, and providing career information and skills. The cognitive domain lies at the intersection of education and psychology, with members of both disciplines contributing major typologies of levels of learning.

Classification 1

Stated Goals for Undergraduate Instruction in Sociology[a]

1. To transmit a body of knowledge to the students.

    Concepts, theories, empirical findings, research methodology, alternative perspectives, substantive areas, the nature of the discipline and its overall perspective.

2. To develop certain substantive understandings in the students.

    Understanding of everyday life, culture, and the social forces affecting the life-situation of self and others.

    Practical insights into social problems and issues, in preparation for playing a responsible part in society.

3. To contribute to the general intellectual and personal development of the students.

    Contributions to general or liberal education.

    Development of critical thinking and analytical skills.

    Personal growth and self-understanding.

4. To contribute to students' vocational preparation.

    Career skills and information about possibilities for future employment.

    Preprofessional preparation for graduate school.

[a] Adapted from Bassis, Goldsmid, Perrucci, Stauffer, and Vaughan (1978). This summary statement was developed from goals statements found in a nonrandom sample of current textbooks, study guides, and instructor's manuals for introductory sociology; course syllabi for introductory sociology collected by the ASA Undergraduate Education Section; some SubTask Group Reports in the same project; and surveys of departmental chairpersons regarding the goals of their departments. The goals are in rank order, with Goal 1 receiving overwhelming support from all sources and Goal 4 receiving fairly limited support. Support for the first statement under Goal 4 is very recent, as the preceding chapter by Perrucci et al. implies.

In the field of education, the first fruit of the University Examiners' attempt to identify educational objectives was the Bloom et al. (1956) taxonomy of the cognitive domain. The taxonomy, reproduced below, is divided into two parts. The first part, called Knowledge, contains only one category, while the second part, Intellectual Abilities and Skills, contains the remaining categories of comprehension, application, analysis, systhesis, and evaluation.

The authors define knowledge as "those behaviors and test situations which emphasize the remembering, either by recognition or recall, of ideas, material, or phenomena" (Bloom et al., 1956: 2). The student simply can reproduce the learned item. However, judging and relating may be involved to the extent that the test situation differs from the original situation. The authors point out that knowledge is the most common educational objective and make a case for its importance. However, they also point out that few teachers consider it the primary outcome of instruction, also wanting students to be able to use their knowledge, to apply it to new situations and problems. Their term "intellectual abilities and skills" refers to these more complex cognitions, often referred to by others as "critical thinking," "reflective thinking," or "problem-solving."

The Bloom typology is designed to be useful to educators, providing examples of educational objectives and illustrative test items at each level. The authors consider the typology as a hierarchy in which each level subsumes the skills and knowledge in the levels of less complexity. This idea was substantially supported in an empirical study by Kropp and Stoker (1966). However, a re-analysis of their data (Madaus et al., 1973) suggests that a general intelligence factor operates, especially at the higher levels.

While the major stated goal for sociology instruction--transmitting a body of knowledge--lies at the first level of the Bloom typology, it is a fair assumption that we want students to be able to do something with this knowledge beyond recall. Our other stated goals call for developing understanding, insight, critical thinking, skills, and preprofessional preparation. Clearly, we are interested in developing the higher cognitions in our students. The Bloom typology can help us know what they look like.

Classification 2

The Cognitive Domain: Bloom et al.[a]

Level		Description: characteristic behavior
1. Knowledge		recalling specifics and ways of dealing with specifics, remembering universals, abstractions, generalizations, and theories.
2. Intellectual Abilities and Skills		
	i. Comprehension	understanding material so that one can translate it, interpret it, or extrapolate from it.
	ii. Application	applying abstractions (ideas, rules, principles, theories) in particular situations.
	iii. Analysis	identifying elements and parts, relationships, and organizational principles.
	iv. Synthesis	forming wholes from elements and parts, assembling them into a new pattern.
	v. Evaluation	judging the value of material and methods in relation to a given purpose.

[a] Adapted from Bloom et al., 1956; Bloom, Hastings, and Madaus, 1971.

Let us take the concept, social system, and see what student behavior might be expected at each level of the Bloom typology. At the knowledge level, students might be expected to recall a definition of the term. At the comprehension level, we might give students a technical definition of the term and expect them to restate it in their own words. In application we might expect students to show how the concept fits or does not fit some organization such as a prison or a school. We might call for analysis by asking students to trace the relations of parts in some complex social system described in an assigned reading. As for synthesis, we might give students descriptions of two very different social systems and ask them to generalize from these as to characteristics of any system. We might ask students for their evaluation of the utility of the system concept in understanding social organization.

Using the Bloom categories, we can make decisions as to the desired levels of learning for each component in our courses. For some elements, the knowledge level may suffice, while for others the higher levels may be called for. We can also use the taxonomy to insure that we have included all levels in our courses, if such is our wish. For example, the objectives for introductory sociology in the writer's Department, presented in Classification 3, range from knowledge of selected concepts to evaluation of sociological analysis itself, with stress on the upper levels of cognition.[5]

In the field of psychology, a major taxonomy of learning was developed by Gagné (1966, 1970), identifying the eight types in Classification 4. In the first type, signal learning, the responses are general and emotionally suffused, such as associating a shout from one's parent with future punishment in a fear response. The next three types include more precise stimulus-response learning and development of chains of stimulus-response learnings, both motor and verbal. The fifth type involves differentiating among similar stimuli, while the sixth type is finding the commonality among stimuli that appear different (concept learning). The higher levels of concept learning and the last two types, rule learning and problem solving, represent the more complex intellectual abilities.

Although the typology includes emotional learning and motor learning, Gagné sees it as lying within the cognitive domain. He says that "its scope is restricted to what may be termed the intellectual, or subject matter content that leads to improvement in human performances . . . " (Gagné, 1970: 25). It seems especially useful in calling our attention to the nature of learning underlying and necessary to the more complex cognitions.[6]

Also in the field of psychology Piaget (1969) has given us a developmental model of thinking, including the sensorimotor stage, development of semiotic functions, development of concrete operations (relating directly to objects), and formal operational thinking dealing with abstractions. While we assume that our students have developed formal operational thinking skills, this may or may not be so. In fact, it has been suggested that only forty percent of high school graduates have this capacity (Whimbey and Barbarena, 1977: 26).

Classification 4

Types of Learning: Gagné[a]

Level	Description: characteristic behavior
1. Signal Learning	making a general response to a signal (conditioned response).
2. Stimulus-Response Learning	making a precise response to a precise stimulus (discriminated operant conditioning).
3. Motor Chaining	undertaking a sequence of two or more stimulus-response behaviors.
4. Verbal Association	making a verbal response to a verbal stimulus in a sequence.
5. Discrimination Learning --Among Previously Learned Parts	differentiating among similar stimuli, with differentiated responses to them.
6. Concept Learning	making a common response to stimuli that appear very different but have some underlying commonality.
7. Rule Learning	responding to a class of stimulus situations in terms of rules (relations between concepts) related to that class.
8. Problem Solving	using concepts and rules in generating new concepts and rules to define and solve a problem.

[a] Adapted from Gagné, 1970; Snelbecker, 1974.

There have been some college programs designed to teach formal operations, with some success (Tomlinson-Keasy and Eisert, 1977; Whimbey and Barbarena, 1977). There has also been an interesting study on the extent to which sociology instruction teaches "critical thinking" (Logan, 1976). It seems that instruction in sociology does not teach critical thinking unless it is an explicit and worked-toward course objective. Typologies in the cognitive domain can help us build objectives at the higher levels and develop strategies to meet them, since we can tell what they look like.

The Affective Domain

Teachers often recognize links between cognitive and affective domains, believing that interested students learn more--and that students with knowledge about something become more interested in it. The authors of the major taxonomy in the affective domain state that there is no fundamental separation between these domains, that the distinction is basically analytical (Krathwohl et al., 1963: 45).

This major taxonomy was developed by the University Examiners group, using the organizing principle of internalization. The authors treat the various levels, shown in Classification 5, as parts of a continuum. They say (Krathwohl et al., 1963: 27):

> . . . the continuum progressed from a level at which the individual is merely *aware* of a phenomenon, being *able to perceive it*. At the next level he is *willing to attend* to phenomena. At the next level he *responds* to the phenomena with a *positive feeling*. Eventually he may feel strongly enough to *go out of his way* to respond. At some point in the process he conceptualizes his behavior and feelings and *organizes* these conceptualizations into a structure. This structure grows in complexity as *it becomes his life outlook*.

To the University Examiners, a major purpose of the cognitive typology was to demonstrate our stress on knowledge, with resulting neglect of higher levels. Their purpose for the affective typology is "to highlight the current emphasis on cognitive objectives at the expense of the affective" (Krathwohl et al., 1963: 57). They discuss the relationship between their two typologies and show that at all levels cognitive objectives have affective components while affective objectives have cognitive components. They discuss using each type as a means of reaching the other. As with the cognitive typology, they include many examples and representative test items.

Classification 5

The Affective Domain[a]

Level	Description: characteristic behavior
1. Receiving (Attending)	being sensitized to and willing to attend to a phenomenon.
2. Responding	actively attending to a phenomenon, being involved with some small commitment.
3. Valuing	internalizing a criterion of worth, being committed to an underlying value.
4. Organization	organizing values into a system, establishing a hierarchy of interrelated values
5. Characterization by a Value or Value Complex	internalizing values to the extent that one is characterized by them, integrating values into a total world view.

[a] Adapted from Krathwohl et al., 1964; Bloom, Hastings, and Madaus, 1971.

The authors point out that teachers who sometimes assume cognitive learning will also develop appropriate affective learnings. In fact, some cognitive learning may develop aversion to the subject. They suggest that learning experiences explicitly directed toward affective objectives are necessary to develop the ones we want. As sociology teachers we seem to want motivated students with interest in our courses and some commitment to them. We may want our students to value the sociological perspective or have positive feelings about what happens in our classes and the information they receive. It is in developing specific objectives for learnings such as these that the affective typology offers direction.

Suppose we want students to have a receptive attitude toward the major sociological insight that social factors affect individual behavior. What might that attitude look like at each level of taxonomy? At the *receiving* level, students might be conscious of social factors that affect individual behavior. (The authors view awareness as very close to a cognitive behavior but involving consciousness of something or taking something into account rather than engaging in recall.) At the *responding* level, students might spontaneously identify social factors that are influencing some particular individual's behavior. At the *valuing* level, students might voluntarily express to others in an informal setting the utility of the sociological perspective on the relationship between social factors and individual behavior. At the level of *organization*, students might assess the behavior of their friends and relatives in terms of the social factors impinging upon them. At the *characterization by a value level*, they might assess problems and social issues in terms of this same perspective. These behaviors are reflections of the desired attitude.

It is more difficult to develop ways to know when these types of objectives are reached than it is for cognitive objectives. However, the examples and sample test items provided by the authors can be helpful. In addition, there is available a small paperback (Mager, 1968) that is very useful in dealing with this difficult domain.

Other domains: The Experiential and Interpersonal

Some teachers may be uncomfortable with the idea that cognitions, affects, and psychomotor responses can be separated, even for analytical purposes. Steinaker and Bell (197[illegible]), point out that the taxonomies we have considered do not treat human experience as a totality. They state (p. 14):

> No single taxonomy exists for looking at experience as a gestalt human activity. Reflecting on our own past we can quickly discern that human experience is not just cognitive, affective and psychomotor; indeed it is all three--and more.

They build a taxonomy of human experience with the levels of exposure to an experience, including responding to it and accepting it; participation in an experience; identification with an experience and positive sharing of it; internalization of experience with changing attitudes and life-style; and dissemination of experience to others, stimulating them to have the same experience. The authors note that experience can terminate at any level and that, while their categories are couched in positive terms, experience can be either positive or negative. In discussing the relation of their typology to those of the University Examiners, they feel that it is more closely related to the affective than to the cognitive. However, they note that the levels of their taxonomy indicate the sequence of activities needed to bring learners up to the higher levels of cognition.

It is interesting to note that the steps in the experiential typology are probably those through which we went in the process of becoming professional sociologists. The typology provides ideas about sequencing learning experiences, although it is not spelled out with extensive examples as are those of the University Examiners.

Among the stated goals for sociology instruction, we sometimes find the goal of contributing to the personal growth and self understanding of our students. The Menges and McGaghie (1974) typology deals with outcomes when the group setting itself is used to promote these learnings.

The levels in their typology include engagement--opening oneself to others; particularity--considering nonintellectualized data; authenticity --accurately expressing and receiving emotions; collaborative decision-making; expanded choice of interpersonal behavior with increased options and feelings of freedom; and synchronism--learning simultaneously with the whole group. They state that, while interpersonal learnings have cognitive, affective, and psychomotor aspects, they have these other characteristics that arise only in an interpersonal situation.

They suggest that their typology may be useful in stimulating further analyses of interpersonal learning and in helping to plan intervention strategies designed to bring about these learnings. As with the experiential taxonomy, this typology lacks extensive exemplification.

Conclusion

The problem of establishing course goals that express what we really want to accomplish through our teaching requires serious consideration and considerable self-examination. However, the fruits are worth the effort, in increased ability to plan and carry out strategies to meet our goals and in increased ability to evaluate our students and our courses. The process of establishing our goals can be helped by making use of the taxonomies in the various domains of learning.

The major virtues of the typologies we have considered lie in (1) guiding our selection of the level of learning we desire for the various components of our courses and (2) insuring that we consider for our courses the variety of types and levels of learning. We are reminded that there are mental abilities higher than recall and are aided in building objectives suitable to them. We are also reminded that cognitions involve affects and affects involve cognitions. We are alerted to the experiential and interpersonal dimensions of learning.

The work of developing concrete application and exemplification of the various typologies in sociology is just beginning.[7] As this continues, we need to share with each other our experiences with the typologies and our attempts to develop in our students the kinds and levels of learning that we desire.

Notes

*My thanks to Richard J. Peterson for his comments on an earlier draft of this chapter and to Michael S. Bassis, Charles A. Goldsmid, and Robert E. Stauffer for their help in assessing stated goals for undergraduate instruction in sociology.

1. In the literature on instructional objectives, the distinction between goals and objectives is based on degree of specificity. Goals are relatively imprecise statements of instructional intent, while objectives are more precise, containing a behavioral indicator of the desired outcome. For example, if one of our course goals is to promote understanding of major concepts used in social research (i.e., variables, hypotheses, and association), we might specify the acceptable end performance as the ability to identify exemplars of these concepts in a given research study (Vaughan and Peterson, 1975). The goal of "promoting understanding" is not clearly recognizable, while the objective of developing the ability to identify exemplars of research concepts is.

For some general references on specifying instructional objectives, most of which include reviews of the debate over their use, see McAshan (1970), Mager (1962), Cohen (1970), and Kibler et al. (1974). References dealing with specifying instructional objectives in sociology include Goldsmid (1979), Goldsmid and Wilson (1980), and Vaughan and Peterson (1975).

2. The perceptual domain has also been defined as awareness of entity and relations of parts, insight, and problem-solving (Moore, 1967). This definition appears to be very closely related to the cognitive domain.

3. Readers interested in levels of learning in the psychomotor domain should see Simpson (1966), also summarized in Kibler et al. (1974) and Bloom et al. (1971), and Harrow (1972).

4. Some efforts have been made at developing multi-dimensional taxonomic models. For example, Guilford (1959, 1967) develops a three-dimensional 120-cell model of the intellect, while Tuckman (1972) builds a 48-cell "Domain-Process-Object" model. These models are overly complicated for our purposes.

5. Sanders (1966) has adapted the Bloom et al. typology to produce a classification of types of questions to be used in the classroom. He defines critical thinking as "all thought processes beyond the memory category"

(p. 6) and considers the main purpose of the classification as insuring use of questions at the higher levels.

6. While the Gagné typology has been developed, modified, and used by many others, there are some modifications of special interest in showing the relationships between psychomotor, affective and cognitive learnings. They include Merrill (1971) and Tiemann and Markle (1973).

7. For an attempt to apply the Bloom et al. cognitive taxonomy in developing instructional objectives for teaching economics, see the Appendix in Ryba and Drake (1974).

References

BLOOM, B., ENGLEHART, M., FURST, E., HILL, W. and D. KRATHWOHL (1956) Taxonomy of Educational Objectives, Handbook I, Cognitive Domain. New York: David McKay Co.

BLOOM, B., HASTINGS, J. and G. MADAUS (1971) Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw Hill Book Co.

BRADSHAW, T. and S. MCPHERRON (1977) "Data report: issues and resources in the undergraduate sociology curriculum." (unpublished)

COHEN, A. (1970) Objectives for College Courses. Beverly Hills, California: Glencoe Press.

GAGNE, R. (1966, 2nd ed., 1970) The Conditions of Learning. New York: Holt, Rinehart and Winston, Inc.

GOLDSMID, C. (1979) "Why formalize the aims of instruction?" Southfield, Michigan: Presented at the KAKN Proceedings.

GOLDSMID, C. and E. K. WILSON (1980) Forthcoming book on teaching sociology. Belmont, California: Wadsworth Publishing Company.

GUILFORD, J. (1959) "Three faces of intellect." American Psychologist 14: 469-479.

--- (1967) The Nature of Human Intelligence. New York: McGraw-Hill, Inc.

HARROW, A. (1972) A Taxonomy of the Psychomotor Domain. New York: David McKay Company.

KIBLER, R., GEGALA, D., BARKER, L. and D MILES (1974) Objectives for Instruction and Evaluation. Boston: Allyn and Bacon, Inc.

KRATHWOHL, D., BLOOM, B. and B. MASIA (1964) Taxonomy of Educational Objectives, Handbook II, Affective Domain. New York: David McKay Company.

KROPP, R. and H. STOKER (1966) "The construction and validation of tests of the cognitive processes as described in the taxonomy of educational objectives." Florida State University.

LOGAN, C. (1976) "Do sociologists teach students to think more critically?" Teaching Sociology 4: 29-48.

MADAUS, G., WOODS, E. and R. NUTTALL (1973) "A causal model analysis of Bloom's taxonomy." American Educational Research Journal 10: 253-262.

MAGER, R. (1968) Developing Attitude Toward Learning. Belmont, California: Fearon Publishers.

--- (1962) Preparing Instructional Objectives. Palo Alto, California: Fearon Publishers.

MCASHAN, H. (1970) Writing Behavioral Objectives. New York: Harper and Row.

MCFAUL, T., BISHOP, P., SINGH, S., JACKSON, D., SICARD, J. and D. ROSE (1977) "A strategy for curriculum development: the constructive use of goals and objectives." Teaching Sociology 5: 65-86.

MENGES, R. and W. MCGAGHIE (1974) "Learning in group settings: toward a classification of outcomes." Educational Technology 14: 56-60.

MERRILL, M. (1971) "Classes of instructional outcomes," pp. 173-184 in M. Merrill (ed.) Instructional Design: Readings. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

MOORE, M. (1967) "A proposed taxonomy of the perceptual domain and some suggested applications." Educational Testing Service.

PIAGET, J. and B. INHELDER (1969) The Psychology of the Child. New York: Basic Books, Inc.

REID, S. and A. BATES (1971) "Undergraduate sociology programs in accredited colleges and universities." American Sociologist 6: 165-175.

RUGGIERO, J., SELENGUT, C., SHOSTAK, A., WESTON, L. and S. YELLIN (1977) "The first course in sociology." (unpublished)

RYBA, R. and K. DRAKE (1974) "Toward a taxonomy of educational objectives for economics," pp. 1-40 in D. Whitehead (ed.) Curriculum Development in Economics. London: Heinemann.

SANDERS, N. (1966) Classroom Questions: What Kinds? New York: Harper and Row.

SIMPSON, E. (1966) "The classification of educational objectives: psychomotor domain." Illinois Teacher of Home Economics 10: 110-144.

SNELBECKER, G. (1974) Learning Theory, Instructional Theory, and Psychoeducational Design. New York: McGraw-Hill.

STEINAKER, N. and M. BELL (1975) "A proposed taxonomy of educational objectives: the experiential domain." Educational Technology 15: 14-16.

TIEMANN, P. and S. MARKLE (1973) "Remodeling a model: an elaborated hierarchy of types of learning." Educational Psychologist 10: 147-158.

TOMLINSON-KEASEY, C. and D. EISERT (1977) "Doing and thinking: their relationship in college freshmen." San Francisco, California: Presented at the 85th annual American Psychological Association.

TUCKMAN, B. (1972) "A four-domain taxonomy for classifying educational tasks and objectives." Educational Technology 12: 36-38.

VAUGHAN, C. and R. PETERSON (1975) "Introductory sociology: a behavioral objectives approach." Teaching Sociology 3: 6-20.

WHIMBEY, A. and C. BARBERENA (1977) "A cognitive-skills approach to the disciplines; intelligence beyond Arthur Jensen." CUE Project Technical Paper #2, Bowling Green State University.

Why Formalize the Aims of Instruction?

Charles A. Goldsmid
Oberlin College

As teachers of sociology we proclaim a wide range of goals. We want our students to know more, and more reliably, about the social world and the tools sociology provides to achieve such knowledge. We want them to use sociology to understand their lives and the society in which they live, to develop skills in analysis and critical thinking, to achieve practical insights into social problems.

Such goal statements are seriously meant, suggestive--and obscure. It is hard to tell exactly what they mean. It is doubtful that students will gain a clear view of their destinations or that, so vaguely put, such statements can guide the instructor in designing a course, or providing a benchmark for evaluating the students' or the instructor's work.

Our goals or aims can be transformed into more precise and measurable statements, often called objectives. What is to be gained by such a transformation? Will students learn more, or differently, or more efficiently? Will the instructor benefit? The department? And are there drawbacks in specifying our aims? Are the resulting objectives likely to be trivial? Does their use preclude discovery of the novel? or prevent the teacher from capitalizing on unexpected classroom opportunities?

Before turning to these and related questions, we examine how a teacher can formalize the ends of instruction.

Operational Statements of Goals: The Case of Prof. Jones[1]

Consider Prof. Jones who teaches a course in race and ethnic relations. One crucial concept employed in the course is marginality. Jones has certain goals for her students; she wants them:

> to understand what sociologists mean by marginality and also to have an analytic grasp of the concept and see that it has wide application in all sorts of social settings (not just for members of racial minorities).

Now what does it mean to understand marginality? How would we know when a student understood the term? What constitutes "an analytic grasp"? How would the instructor know when students can see "wide application"? The initial goal statement suggests worthy intentions which, we suspect, would be shared by many sociologists. But we doubt that such goal statements mean much to students. We know their uncertainty when teachers state similar goals for a term paper or essay exam. ("Remember, I want your papers to be analytic, not just descriptive.") Such vague goal statements cannot be very helpful to teachers as they design their courses, devise ways to assess student learning, to react to students' work and to measure the success of their own efforts.

Elimination of goals is not the answer. Goals provide a direction for the teacher. But unless goal statements can be transformed into objectives, clear statements which describe measurable outcomes, teacher and student will be disadvantaged: less learning will take place.

What would a student have to do to show the instructor that he understood the concept of marginality? With the answer to that question, the teacher has clues as to what must be done to enable the student to master the idea and demonstrate his understanding. Two conventional answers to that question (two objectives) are the following: to demonstrate understanding of the concept, marginality, the student will be able to:

1. write from memory, the definition of marginality provided by the instructor; and

2. recognize the definition when it is presented along with several erroneous or less acceptable definitions.

Now we have transformed the general goal statement into an operational one. We know now what the student is to learn, to do. But some readers will dislike our operationalization. We hear objections as this: "Since when is memorization or passive recall of a term understanding? That may be useful, but it's not what I mean when I say that understand. I mean more than that."

And so do we. But note that our hypothetical reader could evaluate our statement. He knew what we meant and could call it inadequate. Because our operationalization had clear meaning, referred to specific behaviors, he could agree or disagree. In contrast, we doubt that many sociologists could have disagreed with the initical goal statement, "to understand marginality." But more important, we doubt that there could be a productive discussion about the goal without putting it with the clarity of an operational statement.

If understanding marginality is more than memorizing a definition, what in fact is it? What would a student do to show he understood? Consider these possibilities: The student will be able to:

3. write a definition of marginality using words other than those used by the instructor or textbook (i.e., he will rephrase); or

4. provide correct examples of marginality other than those used in class (i.e., he will illustrate).

Thus, the student is said to "understand the concept marginality" when he can rephrase it (objective #3), and illustrate it (#4), not simply recall and repeat its definition (#1) or be able to recognize it (#2).

However the reader may define "understanding" we hope he or she will use terms as clear and measurable as those used in these examples.

Now beyond understanding, Prof. Jones' goal was for students "to have an analytic grasp of the concept of marginality and to see that it has wide application...." Analysis requires that we shake out the parts or elements of some whole. When we do so we often find that the same elements are found in other settings. And we would expect to find the marginality-centrality dimension of social structure present in many settings. Jones wants her students to understand this generic quality and its use in analyzing social structure. Her class is reading a text on race and ethnic relations which discusses only the marginality of black Americans. But the concept is useful in understanding other minority groups, including the Japanese in the United States, Jews in Europe, the United States, and North Africa; and it is useful in understanding occupational roles such as chiropractors, pharmacists, and clergywomen.[2] Like other sociologists Jones sees connections between marginality and status inconsistency, role incompatibility or conflict, partial institutionalization and reference group theory.[3] And she knows that various sociologists use marginality to characterize personality, status, or social and cultural groups. In sum, when Prof. Jones says she wants her students to have an analytic grasp of marginality and to see its wide application, these are the sorts of things she is thinking about.

But this list does not offer clear direction, either for students or instructor. We need an operational statement: what might the student do to show that he could use the notion of marginality analytically? One answer goes like this: the student should be able to do in class discussion, on papers and tests the sorts of things the teacher does in class. The teacher defines and explains; and the student does the same. Yet Jones wants students to do more than repeat what she does in class; she wants them to use the concept marginality in original ways, to build intellectual bridges to groups and situations not treated in class. How can she convey this goal (and plan for measuring its achievement)? Operational statements of "having an analytic grasp" and "seeing wide application" will help. For example:

5. Given a social situation not discussed in class, but one with which he is familiar, the student will state whether any category is marginal and, if so, what traits make for marginality (i.e., the student will demonstrate an ability to generalize); or

6. Given a detailed description of a social setting, the student will (a) identify the components of marginality present or absent, (b) compare the degree to which marginality is present or absent in the setting described in contrast to another of the student's choosing, and (c) list for both cases, the probable consequences of marginality found (i.e., the student will generalize, apply, compare, and evaluate); or

7. Asked to explore the possibility of marginal categories among class members, the student (a) will identify at least three of the following possibilities [we assume a hetergeneous class]: women, blacks, other minority groups, the very poor, the very rich, foreign students; (b) will identify the traits that make for marginality; and (c) will specify at least one advantage and disadvantage of marginality, both for the group and for those in marginal positions.

Thus, "analytic grasp" and "wide application" are demonstrated by the ability to generalize to new cases, apply definitions, compare cases and evaluate them in terms of some external standard.

Instructors cannot safely assume that statements such as "ability to generalize" are automatically clear to students. Such clarity comes only when the instructor helps students put the words to work. If Jones wants to extend the notion of marginality to new groups and settings, i.e., to generalize, then she had better show students how to do it. For doing something is the only verifiable measure of achievement (of learning, in this case) we have. Doing generalizing entails: (1) identifying a class of persons who are marginal; (2) analyzing the traits that confer marginality--strangeness, transiency, lack of common skills (including language), powerlessness, stereotypic labels of inferiority; and (3) locating other classes of persons exhibiting such traits.

Many instructors will wish to cultivate the habit of extending ideas to one's own personal experience (for them, the use of sociology to understand personal experience is an instructional goal). Have any of the students themselves been in a marginal position? (With what consequences?) First generation college students? the minority group member? the military veteran? or the older-than-average college student?

Putting goals in operational terms not only clarifies matters for the student and instructor. It clarifies the measure of achievement. Nor is that all. Being a statement of behavior, the operationally stated goal also points to the instructor's behavior: the teaching procedures needed to exemplify the ways that students will ultimately show their mastery.

Learning Objectives: Definition and Basic Issues

Objectives are operational statements of instructional goals which clearly specify the outcomes desired. Objectives speak not to teaching or studying but to their intended results. They employ action verbs which specify, as unambiguously as possible, what the student should be able to do after some unit of instruction and study. They may include a statement of the learning or instructional process ("having read Davis and Moore, the student....") but focus on the outcome of the process. Ambiguous verbs include: understand, know, appreciate, and see. They refer to an outcome but neither convey clear meaning nor specify an observable student behavior. Verbs and phrases such as write, tell or discuss refer to observable acts but not very clearly. More useful verbs include select, calculate, compare, translate, discriminate, and recognize.

When a student can discriminate in writing between A and B we can make an inference about his learning. Learning is what has occured when one can do something later that he could not do earlier. (Teaching is purposive behavior designed to bring about learning.) In common parlance, we say that someone has learned, when they know more. Thus, learning is an inference we make about increased capacity or ability. But unless we can measure some change in performance we cannot assert that someone knows more or has learned. Hence, definitions of learning rest on the notion of measurable change. This requires that teachers be able to specify their instructional goals in operational form: that they use objectives.

The process of specifying goals in operational form helps teachers as they seek to categorize the outcomes of learning. Prof. Jones wants her students to know what marginality is, to understand the concept, and be able to analyze it and apply it. This goal statement contains four verbs denoting different types of cognitive learnings. Only as Prof. Jones specifies the four categories do they take on meaning. What is the difference between knowing and understanding? How will a student know when he understands? Translation of the four goals into objectives helps the teacher answer such questions. Prof. Jones may use a translation such as this:

Verb denoting goal	Verb denoting objective
knows	defines, recalls, recognizes
understands	translates, explains (using original examples)

Several students of learning (e.g., Bloom, Gagné) have developed classifications of cognitive and other types of learnings. These are useful in helping teachers sort out their instructional goals and specify them in more precise language. (This work is the subject of Prof. Charlotte Vaughan's chapter in the KAKN Proceedings.)

Such classifications are also useful in helping teachers enhance their classroom procedures. Consider, for example, the questions teachers ask students in class discussion sessions. Many questions tap memory: we ask who said what? What is the methodology of a particular study? What percentage of a population group has a given attribute? Such questions tap recall and recognition of information. And the information may be worth having.[4] But if we want students to develop their abilities to understand and use information, we will need to ask different types of questions: those which require students to engage in the appropriate intellectual activities. One way we use information is to apply it to new populations or settings. If this is one of the teacher's goals, her questions (as well as her examples in lecture) should call upon (or illustrate) such extension of information. Specifying our goals reminds us of the need to make our instruction appropriate to our goals. It also allows us, and our students, to assess whether learning has occurred. (On the use of different types of questions, see Sanders, 1966.)

There are diverse types of objectives and many ways of using them. They can be used only by teachers--not given to students--as they design a course, plan a lecture or discussion session, select course materials, prepare tests and assignments, and assess student (and their own) performance. Objectives can be prepared for a course, a unit of a course, or a single class. Objectives for a course can be prepared by one instructor, or a group of instructors who regularly teach the course, or by a departmental committee. A comprehensive set of objectives may be prepared or only a set of core or minimum objectives. The appropriate option depends on individual choice and local circumstance.

Potential Benefits of Using Learning Objectives

Four beneficiaries of using objectives can be identified: the student, the course instructor, other teachers, and the department. The benefits do not flow automatically or directly from the statement of the objectives. They depend on the quality of the objectives, how they are used by students as they study and by teachers as they plan and carry out instruction. After discussing some of the potential benefits, we identify some common misuses and problems in the use of objectives.

The ultimate reason to use objectives is that they may enhance student learning.

Students

(1) "One function of presenting behavioral objectives to students may be to provide direction to their learning. By determining exactly what is expected of them, objectives would assist them in discriminating between relevant and incidental or illustrative content" (Duchastel and Merrill, 1973: 63). Thus students are spared the frustrating, anxiety-producing, time-consuming, and unnecessary effort of trying to guess what the teacher expects of them. This may be particularly important in introductory sociology courses since so few of our students have had any prior experience with the discipline and since sociology asks students to examine their experience using new categories and perspectives. (2) Relatedly, objectives can provide organization to the subject matter helping the student integrate the diverse aspects of the course and its materials.

(3) Objectives provide the basis for a better contract between student and teacher. When students are provided with course objectives, before or at the outset of a course, they are able to make an informed choice as to whether to take, or continue in, the course. They know what information skills they can expect to gain as a result of taking the course. (4) And if objectives exist at the outset of a course, it is possible (or easier) for arrangements to be made whereby students can "test out" of a course. If students can meet the objectives of a course without taking it, why should they take it?

(5) When objectives are used, students can have their performance measured against a standard rather than against the performance of their

peers. When standards are explicit and public, all can aim at them, and it is possible for all students (or none) to meet them. When instructional goals are left vague and unstated, teachers usually let class performance set grading norms. (For a useful discussion contrasting absolute and relative approaches to grading, see Michaels, 1976).

The Course Instructor

Stating instructional aims is the first step in selecting the means of instruction (course format, teaching techniques, student assignments, texts and other course materials). It also provides a benchmark for evaluating the student's learning and the instructor success in fostering that learning. It follows that the more specific and clear the statement of aims, the greater the likelihood that the means chosen will be appropriate and effective. And, of course, clearly stated course aims are more valid and useful benchmarks.

There are other benefits to the instructor of using objectives. Their use allows the teacher to make statements of course prerequisites more precise. Not only does this allow students the opportunity to "test out" of a course, but it insures that the teacher is aware of the information and skills he or she assumes the entering student has (and has not).

Explicit objectives aid the teacher in assessing whether the course is meeting its goals (it is hard to know where you are when vaguely stated aims are your guide). Teachers are better able to evaluate and improve their performance.

Other Teachers

Sociology curricula typically include an introductory course which is a prerequisite for subsequent courses. When students have prerequisites, teachers of subsequent courses assume they are ready for the advanced work. For this to work, teachers must know what information and skills students have gained in prior courses. The typical college catalog description or course syllabus does not provide this information except at the most general level. If a curriculum is to include sequencing of courses and offer the opportunity for cumulative learning, teachers must really know what their colleagues are teaching. The need for such information is especially great in departments where, in a short period of time, many different faculty members teach the introductory courses (we know the variety among introductory courses can be substantial). The clear and explicit statements provided by objectives offer one way to share needed information.

The availability of course objectives (perhaps part of a course syllabus; perhaps stored in a departmental file) also aids teachers in their advising roles. Armed with such information, they can better help students match their interests and goals with the content of offered courses.

The Department

If course objectives are available, departments can achieve more rational curriculum development. They can create majors and alternative "tracks" which have little duplication among courses and more cumulative learning. Unless planners know what information and skills are taught in courses, what students can do after taking given courses, an effective curriculum is unlikely. Faculty members will not know whether their intentions for majors, and other categories of students, are actually to be met.

When objectives are available, communication with other departments and professional schools (which require their students to take a sociology course) is more meaningful. And finally, departments whose teachers have course objectives can more adequately meet demands for educational accountability, achieve an effective advising system, and assess instructional and curricular effects (and pursue ways to improve them).

* * * * *

As we noted at the outset, the benefits listed above are potential. They do not automatically flow from the statement of objectives. Other elements of instructional and curricular planning and development must be brought into play. But the clear statement of our aims is surely a first and necessary step.

Arguments Against Objectives

Many criticisms have been leveled against the use of objectives. Some point to genuine shortcomings, others to potential misuse; some raise false issues.

One objection to the use of objectives rests on the difficulty of and effort required in their preparation. It is true that the careful statement of our aims is a demanding task and that, as Cohen (1970: 83-84) points out, "most college instructors are employed, retained, and supported on bases other than their demonstrated ability to define and to cause learning." He goes on to say that "objectives of this type cannot readily be countered except by appealing to the instructors' sense of professionalism." Yet in those departments and schools which do value teaching, the use of objectives offers teachers the opportunity to be more effective and to better demonstrate that effectiveness.

A second demurrer is that use of prespecified goals prevents the teacher from taking advantage of unanticipated instructional opportunities and outcomes. It is probably true that most instruction has unanticipated outcomes (some agreeable to the teacher and student, others not). But as we observed earlier, sets of course objectives can describe only a minimum, not inclusive, set of course goals. The stated objectives need not, perhaps should not, be seen as the basis of 100% of course time. Specification of goals does not, however, preclude the teacher from using to advantage unexpected classroom opportunities. As Popham notes.[5]

> When one specifies explicit ends for an instructional program, there is no necessary implication that the means to achieve those ends are also specified. Serendipity in the classroom is always welcome, but, and here is the important point, it should always be justified in terms of its contribution to the learner's attainment of worthwhile objectives.... the teacher who prefers to probe instructional periphery, just for the sake of spontaneity, should be deterred by the prespecification of explicit goals. [Italics in original]

A third, and frequently heard objection is that "trivial learner behaviors are the easiest to operationalize, hence the really important outcomes of education will be underemphasized." It is true that the simplest aims are the easiest to operationalize. (Most of the objectives found in workbooks for introductory sociology courses are indeed of this type: they ask students to describe, define, identify and list; rarely do the objectives call for students to deduce, analyze, extend or synthesize.) But if teachers have worthwhile goals which are complex, they can work to state them operationally. Whether or not one uses objectives, the tendency to fall back on the convenient and the comfortable is always with us.

The use of objectives can make it easier to attend to important and complex instructional aims. Our case study of Professor Jones demonstrates that complex goals (such as generalization and application) can be usefully stated as objectives. The precision afforded by explicitly stated goals can provide guideposts for our selection of texts and other course readings, design of assignments, assessment of and response to students on their work, preparation of examinations and assessment of instructional work.

Inappropriate or malign use of requirements by administrators that instructors use objectives is a fourth objection. It is noted that in some schools, administrators use objectives as a means of control over the content of courses--a domain which is properly the instructors'. It is of course true that "ambiguous goals and aims have great defensive value" (Cohen, 1970: 86) and that precise statements often make easier targets. But if the objectives describe goals that are worthwhile, having them may make it easier to fight the good fight for them. It may also help teachers and departments defend themselves against false accusations that what is taught is worthless or inappropriate. Where genuine disagreements exist, and where power rests in the hands of others, the problem is real enough. But since statements of the instructor's objectives usually remain within the department, their misuse by others is not a common occurrence.

It is true that measurability implies or facilitates accountability. But accountability itself is not the villain. Teachers should be judged on their ability to bring about desirable changes in learners. The problem is who defines "desirable change."

* * * * *

Each of these and other arguments against using objectives has at least some foundation. There is no instructional procedure without pitfalls and the potential for misuse, and abuse. And these counter-arguments offer, at the least, cautionary notes for teachers considering the use of objectives. But these objections are only partial truths and do not vitiate the importance of clearly stating one's instructional aims.

Research on Objectives

Earlier we suggested that the ultimate rationale for using objectives is that they may enhance student learning. Research findings to date lend only modest support to this rationale. Perhaps this is not surprising when one remembers that objectives are only one tool in the teacher's kit and not a teaching strategy or method. If they are to advance learning, objectives must be integrated into instruction and its planning. As we shall see, there is some indication that when this is done, learning is enhanced.

There are at least fifty empirical studies--none to our knowledge of sociology classes--on whether student learning is helped by student possession of objectives, on specific compared to general objectives, and by teacher possession of objectives. Review and analysis of this research can be found in Duchastel and Merrill (1973), Kibler, et al. (1974: 5-8), and Melton (1978). In general, results of these studies are divided between those finding small, statistically significant helping effects, and those reporting no significant differences.

There is agreement that simply providing students with lists of objectives is not effective. Learning gains will occur only if students actually use the objectives and the gains will be greater when students have had prior experience with objectives or training in their use. See Melton (1978: 293) and Kibler, et al. (1974: 7). Such findings are specially important since in most studies of objectives, students are using them for the first time, typically with little or no training.

There is some evidence suggesting that students who have and use objectives will take less time to learn instructional material than students without objectives. This effect is most pronounced when the students can control the pace of their instruction and learning (as is true in the Keller Plan of instruction and various other self-scheduled instructional formats). Two studies suggest that students "...provided with objectives spend more total study time on the learning task. If we consider the time involved in reading the objectives as negligible, objectives would then seem to increase the amount of attention paid to the instructional materials themselves" (Duchastel and Merrill, 1973: 62).

Some of the conflicting findings on the effects of using objectives may stem from the non-comparability of studies and failure to introduce certain variables. Melton's (1978) review suggests that a variety of student attributes and instructional procedures affect the outcomes of using objectives. Thus, for example, objectives are likely to be ineffective when they are too general or vague, or when students are already highly motivated to learn and conscientious. (This last finding is consistent with research in PSI and TIPS which shows that, while all students benefit from their use, lower ability students--as measured by SAT or ACT scores--benefit most.)

Study objectives are integral parts of several modes of instruction which have been shown to enhance student learning. The clearest example is that of personalized systems of instruction (PSI), also known as the Keller Plan. All major reviews of empirical research on PSI conclude that it is more effective in fostering content learning than conventional instruction; differences between PSI and otherwise similar classes are typically one letter grade. Four empirical studies on PSI courses in sociology reach the same conclusion.[6]

TIPS (Teacher Information Processing System) is a computer-based, rapid-response system which provides students with results on diagnostic quizzes and with learning suggestions based on their quizzes. The system rests in part on teacher specification and use of learning outcomes. Studies of economics and sociology classes show that students in TIPS courses average about one letter grade higher than students in otherwise comparable non-TIPS courses.[7] A third instructional approach, sometimes depending partly on instructor use of objectives, is CAI--computer-assisted instruction. To date, evidence suggests that CAI is as effective as conventional instruction: lecture and discussion. However, some studies, including one of sociology classes, finds that students in conventional classes which are augmented by CAI, learn more than students using either approach alone. (See Goldsmid and Wilson, forthcoming, Chapter XII.)

Earlier we suggested that teacher use of objectives can help improve instruction. That objectives can help teachers select instructional materials and design their courses and specific classes is suggested by Vaughan and Peterson (1975) in their article on introductory sociology. That preparation of objectives can improve the design and use of questions (for teaching and testing) is suggested by Sander's (1966) book Classroom Questions: What Kinds? The use of objectives in developing teaching skills is suggested by the work of Tillman, et al. (1976).

The Literature on Objectives: A Guide to Knowledge Available

Ample resources are available to help teachers interested in exploring the use of learning objectives in their courses. Vaughan and Peterson (1975) have written a very useful, short introduction to objectives and their application to an introductory sociology course. Some sociology textbooks and many student workbooks and instructor's manuals offer lists of general or specific objectives for each section or chapter. The following introductory sociology textbooks, indicated by their author's names, are examples of those with objectives: Light and Keller; Popenoe; Perrucci, Knudsen and Hamby; Dressler and Carns; and Hardest, et al. Objectives can be found in a few textbooks for other courses; e.g., Kohout's Statistics for Social Scientists.

Most writing pertinent to objectives is found in the literature of higher education and education generally (as contrasted to that on the teaching of sociology). A useful overview, in 200 pages, of research, issues, preparation and use of objectives is Kibler, et al's. (1974) Objectives for Instruction and Evaluation. Following is a listing of major issues or topics concerning objectives and citation to literature dealing with them (the bibliography for this chapter contains the full citations).

Writing Objectives	Mager (1975); Cohen (1970: 9-64); Kibler, et al. (1974: 32-62)
Objectives for Different Types of Learnings	Kibler, et al. (1974: Chap. 5 & Appendix C); Gagné (1974: Chap. 4)
Uses of Objectives	Popham (1973: Part Two & passim); Kibler, et al. (1974: Chap. 1, 7 & passim)
Criticisms and Caveats	Cohen (1970: 83-88); Popham (1973: 43-49); Kibler (1974: 9-20)
Specimen Objectives: in Sociology	Vaughan and Peterson (1975) [Course level]
Specimen Objectives: other fields	Cohen (1970: 89-112) and throughout most of books cited in this section
Research on Objectives	Duchastel and Merrill (1973); Melton (1978)
Assessment of Instruction, Grading, and Objectives	Kibler (1974: 114-142); Michaels (1976); Popham (1973: Part Four); Popham and Baker (1970: Chap. 9)
Theories of Learning and Objectives	Gagné (1974: passim); Gagné (1977: Chapters 11-12)
Design of Instruction and Objectives	Gagné (1974: passim); Popham and Baker (1970: Chapter 6 and passim); Kibler, et al. (1974: passim)
Developing Teaching Skills and Objectives	Tillman, et al. (1976)
A Teacher's First Experience Using Objectives (first person report)	Barry (1978)
A Programmed Text to Teach Students How to Use Objectives	Kibler (1974: 174-185)
The "Objectives" Movement in Higher Education: Overview	Popham (1973: Part One)

Knowledge Needed

The rationale for using objectives, and their link to theories of cognitive learning, is well-developed. More empirical research on the use of objectives--by students, and by teachers--is needed, especially at the college level and most certainly in sociology. The groundwork for this research is well-established. We have high quality papers (e.g., Melton, 1978) which review existing research, identify areas of agreement, catalog variables affecting use of objectives, suggest explanations for anomolous findings, and offer proposals for further research.

The state of the art of objectives in sociology is difficult to assess since the literature is so small. Sociologists who, as classroom teachers or textbook authors, have used objectives would do a service to their colleagues if they wrote descriptive accounts of how objectives have shaped their instructional work. We need to know what, if anything, they have done differently because they used objectives. The next step would of course be empirical research assessing the outcomes for students of that work.

Teaching has both generic and discipline-specific aspects.[9] It would be useful to teachers of sociology to have case histories (and later, empirical studies) which discuss special problems and opportunities of translating *sociological* goals into objectives. For example, many sociologists have noted that undergraduate students enter sociology classes

> with a very "psychological" and "individualistic" view of human nature.... They think of individuals in a pre-Cooley and pre-George Mead fashion, as isolated, separate, autonomous beings...essentially independent of society...students have never been taught to see a social relation or a social system.... *students are much more aware of individual variations from some norm than of the norm itself*.... related to their insensitivity to social structure, I find students inclined towards a strictly psychological interpretation of social life and social problems... It takes patient training of this sociological sight to enable them to perceive the invisible social structure. (Komarovsky, 1951: 253-254; italics in original)

Specification of such particularly sociological goals is needed. We have some indication that the payoff from this work is useful. When Logan (1976) translated his critical thinking goals into operational statements he was able to build the goals into his sociology course and assess the degree to which that course and others in his department met the stated goals. When Davis (1978) carefully defined some goals for the learning and analysis of social facts, he was able to design computer-assisted instructional program. And Vaughan and Peterson (1975) have suggested how specification of goals was useful in designing their introductory course.

We should not be satisfied with the objectives appearing in most introductory sociology workbooks. Typically these objectives are adequate only as statements of the information an instructor hopes students will learn. They call only on skills of recall and recognition. Too rarely do objectives tap skills in using information, e.g., analyzing, extending and synthesizing. Hence, our goals that students develop the sociological imagination, or learn to appreciate the power of the sociological perspective remain vague, if worthy statements of our intentions. (As we have seen, this need not be the case.)

We need descriptive and empirical work on how teachers have used objectives in shaping their instruction and how they have helped students use objectives as they study.

The foundation for needed work is available .

Conclusion

Sociologists know that the test of big ideas lies in their empirical trial. To test them, the ideas must be walked down the abstraction ladder until they can be put in operational terms, i.e., until we can operate with them, get indexes of them, measure them, vary them, manipulate them. Ironically, when they teach. sociologists seem about as likely as ordinary mortals to state their goals in useless abstractions. They sound and are virtuous, but lack clear referents. So foggily put, the instructors cannot know where they are going (nor can students). Therefore instructors cannot know what means to employ to get there. Indeed, they cannot tell whether they have arrived. And arrival is the real test: does the student get there? The more effective instructor is the person who can help more students get farther along the road to an identifiable destination. Additionally, without specified learning objectives the department will have difficulty shaping an effective curriculum and rewarding effective teaching.

The resources to formalize the aims of instruction are available. While we need to know more about objectives and consequences of their use, we know enough to clear away much of the fog that so often envelops instruction in our discipline.

Notes

1. This section appears in somewhat different form in Goldsmid and Wilson's forthcoming *Passing on Sociology* (title tentative).

2. For a useful discussion of the marginality of black Americans and native Americans, see Simpson and Yinger (1972: 185-191). Articles in Pavalko (1972: Chapter 2) treat the "professional marginality" problems of "status and identity" of the occupational groups listed.

3. The link between marginality and status inconsistency is discussed in Simpson and Yinger (1972: 186-187). The same book also discusses the debates among sociologists on the utility of the concept marginality (pp. 185-191). The term marginality was first used by Robert E. Park (1928) although earlier discussions of the same phenomena are found in the work of W. E. B. DuBois (in 1897) and in Simmel's 1908 essay, "The Stranger," to which Park refers.

4. On the importance of basic social facts to undergraduate study in sociology, see Davis (1978: esp. 235-242).

5. Popham (1973: 44). The paper from which this quote is drawn, "Probing the Validity of Arguments Against Behavioral Objectives," was originally presented at the annual meeting of the American Educational Research Association (February 7-11, 1968 in Chicago). It has been reprinted in Popham (1973: 43-49); Kibler, et al. (1974: 9-17), and in many other books.

6. Reviews of the empirical literature on PSI are found in Johnson and Ruskin (1977), Robin (1976), and Kulik and Kulik (1979: 78-84). Two of the better PSI studies on sociology classes are Johnson and Walsh (1978) and Rogers, et al. (1977) both of which cite earlier studies in sociology.

7. A description and evaluation of TIPS use in sociology is found in Bassis and Allen (1977) which cites prior research by TIPS-developer Allen C. Kelley.

8. Cohen (1970: 9-64) is a "Programmed Lesson on Objectives Writing."

9. Discussion of several common problems of *sociology* instruction is found in Goldsmid and Wilson (1980, forthcoming: Chapter VIII).

References

BARRY, R. M. (1978) "Clarifying objectives," pp. 8-33 in On College Teaching, Ohmer Milton, ed. San Francisco: Jossey-Bass.

BASSIS, M. S. and J. P. ALLEN (1977) "Teaching introductory sociology with TIPS: an evaluation." Teaching Sociology 4, 2: 141-154.

BLOCK, J. H., ed. (1971) Mastery Learning: Theory & Practice. NY: Holt, Rinehart & Winston.

BLOOM, B. S., ed. (1956) Taxonomy of Educational Objectives: Handbook I: Cognitive Domain. NY: David McKay Co., Inc.

COHEN, A. M. (1970) Objectives for College Courses. Beverly Hills, CA: Glencoe Press.

DAVIS, J. A. (1978) "Teaching social facts with computers." Teaching Sociology 5, 3: 235-258.

DUCHASTEL, P. C. and P. F. MERRILL (1973) "The effects of behavioral objectives on learning: a review of empirical studies." Review of Educational Research 43, 1: 53-69.

GAGNÉ, R. M. (1977) The Conditions of Learning, 3rd ed. NY: Holt, Rinehart & Winston.

GAGNÉ, R. M. (1974) Essentials of Learning for Instruction. Hinsdale, IL: The Dryden Press.

GOLDSMID, C. A. and E. K. WILSON (1980) Passing on Sociology (title tentative). Belmont, CA: Wadsworth Publishing Co., Inc. (forthcoming)

JOHNSON, K. R. and R. S. RUSKIN (1977) Behavioral Instruction: An Evaluative Review. Washington: American Psychological Association.

JOHNSON, M. P. and E. J. WALSH (1978) "Grade inflation or better comprehension: self-scheduled instruction in introductory sociology." Teaching Sociology 5, 4: 363-376.

KIBLER, R. J., D. J. CEGALA, D. T. MILES and L. L. BARKER (1974) Objectives for Instruction and Evaluation. Boston: Allyn and Bacon, Inc.

KOMAROVSKY, M. (1951) "Teaching college sociology." Social Forces 30: 252-256.

KULIK, J. A. and C. C. KULIK (1979) "College teaching," pp. 70-93 in Research on Teaching, Penelope L. Peterson and Herbert J. Walberg, eds. Berkeley, CA: McCutchan Publishing Corp.

LOGAN, C. H. (1976) "Do sociologists teach students to think more critically?" Teaching Sociology 4, 1: 29-48.

MAGER, R. F. (1975) Preparing Instructional Objectives, 2nd ed. Belmont, CA: Fearon Pubs.

McKEACHIE, W. J. (1978) Teaching Tips: A Guidebook for the Beginning College Teacher, 7th ed. Lexington, MA: D. C. Heath.

MELTON, R. F. (1978) "Resolution of conflicting claims concerning the effect of behavioral objectives on student learning." Review of Educational Research 48, 2: 291-302.

MICHAELS, J. W. (1976) "A simple view of the grading issue." Teaching Sociology 3, 2: 198-203.

PARK. R. E. (1928) "Human migration and the marginal man." American Journal of Sociology 33: 881-893.

PAVALKO, R. M., ed. (1972) Sociological Perspectives on Occupations. Itasca, IL: Peacock.

PAYNE, D. A. (1968) The Specification and Measurement of Learning Outcomes. Waltham, MA: Blaisdell Pub. Co.

POPHAM, W. J. and E. L. BAKER (1970) Systematic Instruction. Englewood Cliffs, NJ: Prentice-Hall, Inc.

POPHAM, J. W. (1973) The Uses of Instructional Objectives. Belmont, CA: Fearon Publishers, Inc.

ROPIN, A. L. (1976) "Behavioral instruction in the college classroom." Review of Educational Research 46, 3: 313-354.

ROGERS, S. J., W. A. SATARIANO and E. D. ROGERS (1977) "Self-scheduled introductory sociology: an evaluation." Teaching Sociology 4, 2: 125-140.

SANDERS, N. M. (1966) Classroom Questions: What Kinds? NY: Harper & Row. (extension of Bloom taxonomy)

SIMPSON, G. E. and J. M. YINGER (1972) Racial and Cultural Minorities, 4th ed. NY: Harper and Row.

TILLMAN, M., D. BERSOFF and J. DOLLY (1976) Learning to Teach: A Decision Making System. Lexington, MA: D. C. Heath.

VAUGHAN, C. A. and R. J. PETERSON (1975) "Introductory sociology: a behavioral objectives approach." Teaching Sociology 3, 1: 6-20.

Assessing the Means We Use: Towards a Rational Model of Course Development

John F. Schnabel

West Virginia University

Richard J. Gelles

University of Rhode Island

William D'Antonio

University of Connecticut

Despite what many of our undergraduate students may believe, sociology faculty are not recruited by placing advertisements on the backs of matchbooks and they do not arrive at the front of the classroom without training in the discipline of sociology. The route of professional socialization which most sociologists travel acquaints them with numerous courses in the sociology curriculum, a variety of teaching strategies, and a passing acquaintance with available instructional methods and materials. Most young professionals will get their feet wet by serving as teaching assistants, an exercise which can range from taking attendance and correcting exams, to being responsible for teaching part, or even all, of a sociology course. A few sociologists are exposed to formal graduate training in the professing of sociology (D'Antonio and Lena, 1976; Wilson, 1976). By and large, sociologists arrive at the lecture with some experience in organizing courses, selecting materials and books and creating and grading examinations.

The purpose of this paper is to examine what types of knowledge and resources are available to sociology instructors as they plan their courses. In addition, we also discuss what kinds of knowledge and resources are needed to help instructors prepare their courses and to assist sociologists in organizing departmental teaching programs.

In order to assess the knowledge available and knowledge needed in planning a course, we have, in the best tradition of sociology and Max Weber, prepared a rational model of course development. The fifteen-step model[1] (see the following list) can be viewed as a series of decision points where instructors can, if they so desire, make rational decisions based on resources and knowledge available. We should emphasize that the model is presented for its heuristic value in stimulating a discussion of what knowledge is available and what knowledge is needed by instructors in planning courses. We recognize that in many cases instructors do not acknowledge the steps we have listed as decision points. Some decisions are implicitly made as a consequence of the instructor's experiences and professional training. Other instructors may not view some steps as relevant in preparing to teach a course. Many instructors may make most of the decisions we list without a major analytic investment in the costs and rewards of the decision. Nevertheless, our goal is to assess what assistance is available to instructors who ponder all or some of the 15 decisions we list.

Towards a Rational Model of Course Development

Step 1	Analysis of the institutional context or environment
Step 2	Assessing curricular goals and objectives
Step 3	Assessing how the specific course fits into the curriculum
Step 4	Setting specific course goals and objectives
Step 5	Identifying which students the course is designed for
Step 6	Identifying the instructor's own implicit or explicit model of how students learn
Step 7	Determining the boundaries of the course
Step 8	Determining what the master framework for the course will be--how the course will be organized
Step 9	Deciding how the course will be broken down--modules, units, and so forth
Step 10	Setting the overall presentation strategy--e.g., lecture, lecture-discussion, seminar, and so forth
Step 11	Choosing what materials will be used (e.g., text, visual, audio, simulations)
Step 12	Determining the presentation strategy for each unit
Step 13	Deciding what assignments students will carry out outside the classroom
Step 14	Selecting a method of assessing or evaluating students
Step 15	Selecting a method to evaluate the course

Step 1. We propose that the first step in developing a sociology course is to assess the institutional context in which the course will be taught. The department, college (and university) constitute the teaching environment. The teaching environment provides both the institutional resources and the institutional constraints which impinge on the teaching of the course.

Time, money, and physical space effect how large the class will (or must) be, what resources are available to the instructor (e.g., films, visual aids, assistants, copying funds). There are institutional rules and regulations--course loads, rules about text selection for multiple sections, final examinations, and grading procedures all relate to how a course will be presented. In addition, an instructor will have to assess, at some point, his or her role within the institution. Is time to be divided between teaching, scholarship, and service? Are there major role conflicts which arise? How does the institutional context affect the instructional and professional role? What activities are desired, supported, and rewarded by the institution?

Step 2. The departmental curricular goals and objectives are related to the institutional context of teaching and also require the concern of the sociology instructor. Curricular goals influence who the clientele will be. Is the department focusing its energies on preparing students for graduate work in sociology? Is the main mission of the curriculum to prepare the students for employment? Are the sociology courses in the curriculum mainly there to provide service to other majors and departments in the college or University? Rarely does a department develop a curriculum solely to satisfy one purpose. Thus, instructors tend to have multiple missions and serve multiple clienteles. Before embarking on a course, then, the instructor needs to consider how best to adapt the work to the several missions and differing customers.

Step 3. Having assessed the department curriculum, the next step is for the instructor to examine how the course fits into that curriculum. Is it required for the major? Is the course required by other deparmments or programs? Is the course a terminal point in the curriculum or a prerequisite for other courses?

Step 4. Once an instructor has located a course within the institutional setting, department mission, and curriculum, the next logical step is to set specific course goals and objectives. Such objectives, in the ideal world, should be tied to both the student and the curriculum. Such goals can be cognitive, behavioral, or developmental. (See the preceding chapters.)

Step 5. Although this step is probably related to, and inextricably tied to the previous steps, we chose to separate from the previous steps the assessment of the specific students enrolling in the course. Ideally, course goals and objectives are general in nature and are not tied to individual student traits such as age, gender, or ability. But once such general goals have been set, they must be considered in terms of the actual students. The course goals can be tailored to whom and where the students are. The following steps in our rational model of course development all return to this point.

Step 6. Instructors implicitly or explicitly hold models of how students learn. Such models or assumptions about the learning process affect and often determine the way we will teach a course. Ideally, models of learning will be empirically based in learning theory and what is known about how students learn.

The preceding steps are all important in arriving at decisions on how a course will be presented. For the purpose of this chapter, we will concern ourselves with the nine steps which follow. We consider each step in terms of the knowledge available from the research on teaching and what knowledge we would desire from future research which would assist us through each of these steps.

The steps which most instructors will recognize are the steps in course organization which begin when the instructor outlines the course and/or prepares a course syllabus.

Step 7. The boundaries of a course depend, in large part, on how the course fits into the larger sociology curriculum. Basically, an instructor must assess what theories, concepts, ideas, or empirical evidence will be included in the course and what material shall be excluded. Such decisions depend on curricular goals, course goals, course size, resources available, and the time within which the course is presented. Our concern is what resources are available to help an instructor establish the boundaries of a specific course.

Step 8. Within the boundaries of a course an instructor will typically establish a master framework or course format, the general organization of the course. The decisions made at this point concern the setting in which learning occurs. The master framework is different from the specific techniques an instructor will use to present the course material--in fact, techniques depend on decisions made about the master framework.

In establishing the course format an instructor will make decisions on such issues as: behavioral objectives, themes, theoretical orientations, inductive or deductive approach, chronological approach, sequential approach, inquiry method, problem-solving method, self-paced instruction, and guided design.

The range of options in organizing a master framework may be grouped into five general types:

1. Lecture-discussion
2. Small group instruction
3. Individualized instruction
4. Seminar
5. Experiential--apprenticeships, student research

Step 9. At Step 7, when the instructor establishes the course boundaries, one consideration is the time within which the course is to be taught, a semester, a quarter, a trimester. At Step 9, the temporal contingencies become important when the instructor has to determine how the course will be organized into segments. One way is to identify the natural breaks in the course material. Some instructors do this by focusing on issues or problems, others assess the course text and assign one chapter to be covered per week. Also to be considered are the time blocks available for each class--60 minutes, 90 minutes, or whatever.

Instructors can choose from a number of organizing strategies--modules, PSI units, TIPS, mini courses, programmed learning, or other techniques.

Step 10. At Step 10, the instructor arrives at the point where the question "How shall I teach the course?" is answered. In establishing the overall course presentation strategy, an instructor can select from a number of pedagogical roles--lecturer, discussion leader, manager of materials (as in a PSI course), facilitator, seminar leader, or initiator of learning contracts with students. The course presentation strategy also involves altercasting students into their roles--peers, partners, assistants, presenters, writers, actors, debaters, tutors, listeners, et al.

Step 11. Although selection of materials is Step 11, it is doubtful whether an instructor makes many of the previous decisions without having some idea of what instructional materials will be used in the course. Ideally, however, the course materials are selected after the course goals and objectives have been established, boundaries set, units considered.

With these decisions reached, instructors can consider what they want students to read, see, hear, or experience. The choices are nearly infinite. One can have students read texts, books or readings, newspapers, journals, novels, magazines, case studies, field notes, and computer print-outs. Students can listen to tapes, records, radio, guest speakers, the instructor. Films, videotapes, television, photographs, slides, and filmstrips can be viewed. Games, simulation, internships, and research can be experienced.

With the vast range of choices, what resources are available to help an instructor make decisions?

Step 12. Although the instructor selects an overall presentation strategy at Step 10, that strategy may have to be modified given the nature of the material or unit which is being covered. Specific concepts, ideas, theories, or points of view may require specific instructional approaches.

Step 13. Out-of-class activities largely depend on the role of the student which is indicated at Step 10 and again at Step 12.

Step 14 and Step 15. These last steps require instructors to determine how they will evaluate the students and themselves.

The discussion which follows focuses in summary form on the knowledge available and the knowledge needed to arrive at decisions at selected steps as a way of suggesting that much more is needed than what we presently have.

Step 7 - Determining Course Boundaries

There are no empirical studies nor any anecdotal accounts written about how instructors go about establishing substantive, conceptual, theoretical, or disciplinary boundaries for their sociology courses. We suspect that the modal pattern of arriving at decisions about what to include and exclude in a course, especially for young professionals, is to rely heavily on the courses they have taken as undergraduates or graduate students and the courses for which they have served as teaching assistants. Casual observation has led us to believe that many a new instructor has come to grief trying to use a graduate course outline as the basis for structuring an undergraduate course.

There are no set standards or criteria for what must be included in a specific sociology course. Presumably many courses are structured using a recipe technique which calls for adding one part core concepts, one part major figures in the field, a dash of relevant or trendy issues, and one part issues in which the instructor is interested (or has examined in his or her dissertation or last journal article).

But there are certainly instances in which instructors find themselves suddenly thrust into the breach, having to structure a course that they: never had as an undergraduate or graduate student, never assisted, never examined as a research interest, and/or know little about. What do instructors do in these instances? What resources are available? Who or what can be turned to as a guide to help construct the boundaries for the course?

Course Syllabi

One quick way of assessing and establishing course boundaries is to consult a course syllabus. Course syllabi from previous courses the instructor took as a student may be available. Or, an instructor can go to a department file (if one exists) to check out course outlines used by colleagues. If no file exists, a colleague's outline can be borrowed. The Undergraduate Education Section of the American Sociological Association has published and distributed course outlines for introductory sociology and social problems. (Course outlines for the family course are currently being prepared.) The first issue of the New England Sociologist published a course syllabus for a deviant behavior course (1978).

Texts

Although we listed choosing a text as Step 11 in our rational decision model, an instructor can review textbooks published in a specific content area to see what ideas, issues, and themes are common to most texts (and by inference common to many courses). The texts, usually 20 chapters or more in length, provide a rough boundary; but many instructors find that they still face choices of what chapters (issues) to include and exclude in the final course.

Handbooks

Instructors can turn to a series of handbooks published by Rand McNally for overviews of theories, methods, and issues in specific substantive areas. These handbooks could be used to provide outlines and course boundaries. The available books cover the fields of marriage and the family (Christensen, 1964); social problems (Smigel, 1971); personality (Borgatta and Lambert, 1968); socialization (Goslin, 1969); organizations (March, 1964); work, organization and society (Dubin, 1976); criminology (Glaser, 1971); communications (deSola, Pool et al., 1973); and, small groups (Hare, 1976).

Journal Articles

There have been a number of journal articles published which focus on teaching specific substantive courses in the sociology curriculum. Such articles include Barnett's discussion of sociology of development courses (1974), Boocock's innovative course in urban sociology (1970), Boulding et al. on the teaching of the sociology of world conflicts (1974), Clear, a new approach to social problems (1966), and Fuller, also on teaching social problems (1938-1939). Himes (1952) and Gladden (1953) have discussed teaching family courses. Kolack presented an outline for an ethnic studies course (1975). Riley discussed a course on social change (1940), while Washburn published on teaching a political sociology course (1974).

There are many other examples of sociologists publishing articles on the teaching of specific courses (for further references see Goldsmid, 1976; Goldsmid, 1977; Goldsmid and Davis, 1978, and Goldsmid and Goldsmid, 1979). Each of these articles provides guidelines and suggestions for establishing course boundaries.

Book Review Essays

Occasionally, sociology journals publish book review essays which discuss a number of texts in a specific field. These essays are aimed at helping an instructor make an informed choice about what text to select; but, the essays also review the topics and issues covered in the texts and thus, these essays provide a valuable resource for assessing what is considered important content for sociology courses.

Concepts

Instructors not only have to decide on what substantive issues will be included in a course, but also what concepts are crucial, necessary, or important for building a sociological imagination or sociological perspective in the minds of students. What are the core concepts in sociology and in specific substantive areas? Borgardus published a paper on this issue in 1960, and a more recent review of core concepts in sociology has been prepared by Clarke et al. (1977), members of the ASA Undergraduate Education Projects Task Group on Curriculum.

State of the Art Essays

We once discussed the development of a sociology of sports course with a colleague. Our colleague wanted to design such a course, and felt that it would be a valuable addition to the sociology curriculum. However, he was stumped as to what material to include in the course. We recommended that he consult a review of essays on the sociology of sport which had been published in Sociological Quarterly by Snyder and Spreitzer (1974). This article proved to be quite helpful in setting the course boundaries. The periodic state of the art essay which is published on certain substantive areas can serve as a helpful resource for instructors.

Other Resources

In addition to these materials and resources which are available to help establish course boundaries, instructors can review Sociological Abstracts to see what is published in the area in question. Study guides and instructor guides to textbooks can be reviewed. An instructor can also consult the International Encyclopedia on the Social Sciences.

In short, the process of establishing course boundaries need not be a blind one. One does not have to rely exclusively on old, or borrowed course outlines to construct a new or revised course syllabus. Yet there are resources which are needed to further aid sociology instructors.

Knowledge Needed

While there are numerous resources which are available to aid instructors in establishing course boundaries, there are no available standards or benchmarks which instructors can use to guide their final selections. The instructor is faced with a cafeteria of ideas, concepts, perspectives, issues, and problems. Which ones are essential to teaching sociology? Should an introductory course include a unit on small groups, religion, death and dying? Which social problems belong in a social problems course? What are the core concepts of sociology? Which are the crucial theoretical perspectives? What do students have to know about research methods? Which statistics are indispensable to sociologists?

There are no rigid guidelines for setting boundaries. In fact, it is questionable whether a sociologist even need be concerned with keeping to disciplinary boundaries. Borck (1977) presents the case for and against disciplinary boundaries for sociology in an innovatively written mock trial which lets the reader decide how necessary boundaries are (again the idea is brought up, but no answers provided).

Perhaps, in a discipline which has been characterized as pre-paradigmatic (Kuhn, 1970), it is too early to expect that we can have consensus on what should be in a sociology course or sociology curriculum. It could be that our clientele is so diffused and their needs so widespread that consensus is impossible. If that is the case, then our knowledge needs are for a cafeteria with a greater variety of choices. If this is not the case, then establishing lists of core concepts, theories, and issues is our prime task.

Step 9

On Dividing Classes into Units

There is practically no scientific knowledge available to tell us if it makes any difference in the teaching-learning process whether classes are divided into units, what kind of units, or different length of units.

Most instructors divide their courses into units according to time and topic: that is, they decide how much time they have to achieve their goals, and then divide the topics of the course into time periods. As indicated in the introduction the most common pattern seems to be to use either the chapters of a text or a series of readings (paperbacks, articles) as the basic units. Interspersed may be one or more exams, and occasionally among the more innovative, the use of films or an occasional game or simulation may be attempted. The decision of how much time is to be allotted to particular units seems to be based on the subjective interests of each instructor. Each instructor decides on what kind of units and the length of units based on the perceived importance of particular topics. Frequently, each decides to give a week to each of 12-14 topics around which the course syllabus is built.

This is not to suggest that the decisions are made irrationally. Undoubtedly instructors employ their experience, their sense of the discipline, and an overall view of the particular course to decide the unit questions. Instructors do have some sense of the balance they hope to achieve. They do have a sense of the fact that students learn some topics more quickly than others regardless of the degree of importance of the topic. They know that the breadth vs. depth question depends on their own goals as well as the curriculum goals of the department. The choices they make, therefore, when breaking the course down into units, are what amount to hypotheses about effectiveness and efficiency. But few of us feel comfortable continually hypothesizing. Furthermore, the current state of research in the field is of little help. There is little to help us decide whether twice as much time should be given to social organization as to culture in an introductory course. We can make only rough approximations as to whether we should spend three weeks on probability theory and only one week on chi-square in methods.

The closest approximation to guidelines, for decision-making in breaking the course down into units, has come from ex post facto research designed to assess the strengths and weaknesses of Personalized Systems of Instruction. As far as we know, PSI (which also has other titles) was not developed on the basis of research into the best ways to establish course units even though it is based on assumptions and a working principle of knowledge units. Most of the breakdown into units in PSI has undoubtedly come from the educated hunches of the developers.

To state that, at present, this process is post facto is not a negative criticism, but rather testifies to the fact that some of our colleagues are attempting to be increasingly scientific about sociological instruction and are at least at the point of hypothesis-testing, while most of the rest of us are merely at the point of developing hypotheses.

The key feature of PSI is, in fact, its division of the work into units which are clearly defined. And the key feature of the unit approach is that the student completes each unit at his or her pace. Thus, time is very flexible. The student takes as much time as needed to learn the materials of each unit.

The evaluative evidence to date suggests strongly that students do learn more from PSI than from the more conventional approaches to learning (especially lectures), that they would recommend the course to friends, and that they find PSI a worthwhile intellectual challenge.

Wesley W. Craig et al. (1976) describe a modern systems approach to the classroom ("Creative Learning Through the Application of Sociological Principles," i.e., CLASP). They advocate a small class size with at least one faculty member for every 30 students. The separate units into which the course is divided are discussed by the authors. Pretest and posttest scores are discussed for both experimental and control classes. Personal testimony from students and instructors is offered to show that this can be a valuable innovation. There are more data here than in most articles on innovations.

The utility of CLASP is challenged in part by a subsequent article by Howard Bahr and Bruce Chadwick (1976) who offer a closely reasoned critique. Their major criticism is that CLASP develops an unintended goal of personality development (i.e., development of the whole person, development of personal autonomy) which becomes more important than the learning of sociological concepts and theories. The critique is a valuable addition to the dialogue.

Kay Clifton (1976) argues that the old lecture-test approach is sterile and that "doin soc" offers a more empirically oriented approach to introductory sociology. The evaluation is based only on the author's personal experience.

Van Valey (1976) describes a pedagogical innovation which involves adding a second or mini-course to a regular course. The mini-course expands on some topic or adds topics not covered in the regular course. The possibility is there to bring in new faculty and add depth and new dimensions to a course. No evaluative data are presented.

Other than writings on PSI, the works that mention units at all do so descriptively. A typical example is an article by Kolack (1975) in which she reports on "A Course in Ethnic Studies." She informs the reader that the course was divided into three major time units: the first unit consisting of the first four weeks, background information; next six weeks, data gathering; final five weeks, data analysis and reporting. We receive no feedback on the relative importance of these three units, whether they were of the right time length, of equal value, and so forth. The author says the process worked well and that she would use it again. We find this kind of evaluation in most of the articles which describe teaching innovations. If nothing else, such innovations produce a Hawthorne Effect.

Knowledge needed: Answers to the following questions would help us understand the importance of units to the learning process.

1. Do students learn more in class time units of 50 minutes three times a week, vs. 1 hour and 15 minutes twice a week, vs. 2 hours and 30 minutes once a week?

2. Do instructors teach more effectively in one or the other time units?

3. Are different kinds of courses taught more effectively in different time units? Is 50 minutes the maximum effective time unit for lectures? Might statistics courses be better taught in time periods ranging up to 2 hours?

4. Are goals set for courses realistic in terms of the means (time units) available?

5. Are such innovations as mini-courses and modules successful simply because they are innovative and break the monotony of the classroom routine, or may their success be due in some part to the fact that they use time differently?

6. What do the data from PSI tell us about the time-topic units that other teaching styles use, especially the more traditional style of the lecture format which has a different or continuing topic for each class period, always covering new material to fill pre-ordained course objectives?

7. How do we decide what constitutes a content unit as opposed to a time unit?

8. How do we go about meshing time units and content units?

Step 10 - An Overall Course Presentation Strategy

The amount of knowledge available about the strategy of course presentation contrasts sharply with what is available on breaking the course into units. We have extensive descriptive and evaluative materials on lectures, discussion, and seminar, as well as the new strategy of Personalized System of Instruction.

Much of the research comparing the utility and limitations of lecture vs. lecture-discussion, discussion and seminar approaches to teaching has been well-summarized by McKeachie (1978). Perhaps the most definitive summary of research has been conducted by Bligh (1972), and there is now general agreement that we have about all the research we need on the strengths and limitations of these several strategies.

The Lecture strategy does the following well:

1. presents facts and figures rapidly
2. places the emphasis where the teacher wishes
3. excellent for providing background information
4. can present the same material to any size of class
5. can be filmed or taped for replays as desired
6. can be interesting and lively, like a good show
7. can be broadcast to remote locations
8. teacher controls the content and sequencing of events
9. combines readily with other media such as films, tapes
10. can be interrupted by requests for clarification, more detail

On the other hand, the lecture strategy has the following limitations:

1. it is difficult if not impossible for the instructor to adjust to individual speed of comprehension
2. the communication is mostly one-way
3. there is little opportunity for learner to participate actively and thus to maintain interest

4. there is no direct check on learning taking place
5. it is difficult for the instructor to prepare for an unknown audience
6. it is difficult to maintain attention and interest
7. effectiveness depends on the skill and personality of the instructor
8. learning is largely reduced to what can be tested on a multiple choice examination
9. lecture is much less effective in developing higher cognitive skills, such as the ability to put thoughts into writing
10. especially in large classes, the lecture is not popular with students or instructors

The Discussion strategy does the following well:

1. it encourages learner activity
2. it is easier to maintain student interest
3. the learner can check his or her ideas with others
4. where instructor uses prepared questions, these can stimulate critical thinking
5. learner is given a sense of personal responsibility for learning
6. experiences of advanced students are shared by others
7. it is seen as a mature, cooperative way of learning
8. both instructor and student receive regular feedback on learning and progress
9. with feedback instructor can change pace, be flexible as needed
10. is strongly favored by instructors and students over the more formal lecture strategy

The Discussion strategy has the following limitations:

1. it depends heavily on the learner's capacity and mood for participating
2. it is very time consuming
3. instructor must prepare stimulating questions
4. requires back-up aid if discussion lags
5. requires that the instructor be skilled in human relations, directing and controlling
6. instructor needs to summarize with skill, trying to pull together the main points of the discussion

7. instructor must be able to think fast, to shift, adapt and accommodate
8. requires permissive guidance on the part of the instructor
9. demands that the student read or otherwise be prepared for class in advance
10. class size must be limited to allow for maximum and effective participation over time

(The above summaries adapted from William Freeman).

The Personalized System of Instruction (PSI), also known as auto-instruction, and programmed learning, does the following well:

1. the rate of presentation and learning of materials is under the control of the learner, that is, it self-paces
2. it is specifically designed to meet performance objectives
3. it provides frequent knowledge of results
4. the learner masters one step before the next is presented
5. there is necessarily active learner-participation
6. there is efficient use of time by teacher and by learner
7. the program has been validated at least in some cases
8. the content is standard and reproducible
9. there is stability of presentation from one day to the next

But the PSI is not without its limitations, which are:

1. it takes many hours to prepare and validate each unit
2. much training and experience are needed to produce a particular course for PSI
3. some learners dislike small steps
4. the PSI does not fit into the normal or regular school units of time which are oriented toward the group class approach.

(The above adapted from William Freeman).

The research evidence on PSI has been organized to respond to three main questions:

1. How effective is PSI? In 30 of 31 comparisons between PSI and conventional courses, in which controlled experimental methods were used to test for differences, PSI students performed better, and in 25 of these cases, statistically significantly better.

In six comparable studies, PSI students performed better than students from lecture courses in time intervals ranging from 3 weeks to 15 months.

PSI students performed better in advanced courses in the same field than did students from conventional courses.

PSI students also performed better in related courses than did students who took the same course (biochemistry) conventionally. The PSI and lecture groups were initially identical on such relevant variables as grade point average and medical test scores.

Furthermore, students like PSI; the instructor does not seem to concentrate on just teaching the test, and the positive effects of PSI are broadly evident in related work.

2. What is the heart of PSI? Research evidence supports the conclusion that "the continuous feedback of individual performance toward unit mastery is the most important single feature."

3. Who is the ideal PSI student? In PSI courses evaluated to date, students from each ability level perform at least as well as students in conventional courses. Thus, there is no evidence that PSI is inappropriate for any student or group of students.

(Above from James Kulik, "PSI--A Formative Evaluation," in Memo to the Faculty, the University of Michigan, No. 55, May 1975, p. 5.)

A Note on Student Performance in Different Settings

There are several ways to think about student performance. The most common is simply to combine grades received by students. Using this criterion, we find that PSI tends to yield the highest sets of grades, while the large lecture sections yield the lowest sets of grades. A variety of evaluative tests show that there are no distinctions between grades of students in discussion and lecture courses.

But the simple comparisons may miss some important points about learning. The PSI versus lecture studies show that over time there is more learning by PSI. And when discussion versus lecture approaches are compared, the discussion approach is found to produce, over time, more deeply rooted cognitive learning, more effective critical thinking. Finally, when the discussion approach uses much written work, performance means that students have been disciplined to put sentences together to form coherent paragraphs, rather than merely trying to decipher sentences on multiple choice tests.

Student participation varies by the strategy used. The lecture strategy keeps the students passive, the discussion strategy involves them to varying degrees, and the PSI requires active participation at all stages.

On the Role of the Instructor

In the PSI strategy the role of the instructor is radically altered from what it is in the traditional role of lecturer, and in the conventional role of discussion leader. Essentially, the instructor becomes an organizer of carefully designed materials, a manager who has to oversee a number of student assistants who may range in number up to eight or so. But the instructor also retains the role of teacher, but on a tutorial level, meeting with students who need more help than the assistants are able to provide. And the instructor also works closely with the assistants in what is an important teaching-learning relationship.

A most interesting feature of the PSI is that it usually involves undergraduate students as assistants. These are students who have already taken the course and have shown ability and volunteer. The results of this innovation have been evaluated by a number of instructors, and the evidence is that this has proven to be a very rewarding and intellectual experience for these assistants.

Step 12 - Determining the Presentation Strategy for Each Unit

We spoke to the overall presentation strategies for courses in Step 10. Here the concern is with specialized strategies in the sense that they have been developed for particular segments or particular class days or as specialized supplements. They address the question of what will be done or can be done within the overall presentation strategy to deal with specific time units, content units or in the development of perspectives and/or skills which the instructor believes to be desirable.

While we may have broad objectives for the entire course, as we break it down into units, our goals for particular units may differ from goals for other units. We may be content, for example, that students have committed one thing to memory, while in another instance we might hope that they would be able to generalize or employ specific skills.

The number of strategies developed is quite large and could not be exhaustively listed because they are often individually created, tailored and not reported. But even a minimal listing would include such things as slides, films, videotapes, audiotapes, simulation, games, diaries (journals, ledgers), field trips, role playing, experiments, case studies, models, modules, illustrations, devices, outside resource persons, and computer uses of many kinds.

With but a few exceptions, most of the above have been created to serve a special need which an individual instructor believes to be present in his or her own teaching/learning situation. As a result they are often non-thematic, idiosyncratic, and unevaluated.

Each author tells us what has been done, what the strategy was intended to accomplish, and how well he or she believes it has worked. But seldom have instructors reported any systematic evaluation of the strategy. What little evaluation there has been has usually been achieved by either asking students point blank what they think of the strategy or rendering a judgment based on a non-random sample of students who volunteered their opinions. In short, we are told what the strategy does but without solid evidence to support the claim. The advent of the journal, Teaching Sociology has provided a much needed forum for the exchange of ideas about the tools of our trade. More of these specialized strategies are being shared, and we benefit, no doubt, by getting ideas for our own courses. But one thing that stands out in a review of that material is that heavy reliance is placed on the assumed face validity of using a particular strategy.

In only a few instances is there a body of literature to assist us in making rational decisions about employing particular strategies in accordance with the goals we hope to achieve. Perhaps the most researched of all the strategies has been that of simulations and games. (Cathy Greenblat provides a complete summary of claims and evidence in Teaching Sociology, October 1973.) We now know pretty much what simulations and games can and cannot do--which of our instructional objectives can be reasonably assured by their use and which cannot. A history of the careful and extensive evaluation of simulations and games, however, demonstrates that it came about largely because of conflict. Early proponents were making exaggerated claims for their use which others challenged. Users were virtually forced to respond with evidence.

It undoubtedly overstates the case to suggest that by and large we do not yet know what most of the other specialized strategies really teach or how they teach it. Such an overstatement is intended, not as condemnation, but as a similar challenge to those who know the most about specialized strategies (the users) to begin developing evaluative procedures for those strategies. The content of that challenge carries with it two qualifiers. First, it is a large task and no one of us can do the kind of evaluation that is needed for all of the specialized strategies that we employ. Often we use several of them in a single course.

Second, occasionally a specific strategy may have such a high degree of face validity that formal evaluation might be effort wasted. But more often, evaluation will be useful if not necessary to answer the questions of effectiveness in desired outcomes and confounding side effects.

Given the limitations of where we are at the moment, Step 12 in the proposed model for course development relies heavily on assuming face validity, with all of its virtues and dangers in an effort to fit as clearly and neatly particular strategies with the specific instructional goals of the instructor. In the absence of better evaluative evidence and until it is developed, the clarity of our objectives even to ourselves becomes ever more urgent; else we are risking the waste of valuable instruction time, or worse, sidetracking the learning process.

References

BARNETT, A. S. (1974) "Trends in teaching the sociology of development in the United Kingdom." Teaching Sociology 2 (October): 71-90.

BAHR, H. and B. CHADWICK (1976) "Some comments by fellow travelers: one view of CLASP from the professional peer group." Teaching Sociology 3 (January): 131-137.

BLIGH, D. A. (1972) What s the Use of Lectures? Middlesex, England: Penguin Books, Ltd.

BOOCOCK, S. and E. O. SCHILD (1968) Simulation Games in Learning. Beverly Hills: Sage Publications, Inc.

BOOCOCK, S. (1970) "An innovative course in urban sociology." American Sociologist 5 (February): 38-42.

BORCK, H. (1977) "The case against boundaries." Teaching Sociology 4 (April): 297-314.

BOGARDUS, E. S. (1960) "Selected sociological concepts for beginning students in sociology." Sociology and Social Research 44 (January-February): 200-208.

BORGATTA, E. F. and W. W. LAMBERT [eds.] (1968) Handbook of Personality Theory and Research. Chicago: Rand McNally.

BOULDING, E. M., T. GOERTZEL, J. W. ELDER, and R. H. JACOBS (1974) "Teaching the sociology of world conflicts: a review of the state of the world." American Sociologist 9 (November): 187-193.

BRISK, W. J. (1974) "Student prepared games as a teaching strategy." Teaching Political Science 2 (October): 90-94.

CHRISTENSEN, H. T. [ed.] (1964) Handbook of Marriage and the Family. Chicago: Rand McNally and Co.

CLARK, S. G. and B. GILMAN (1976) Personalized System of Instruction for Sociology: Human Society. Glenview, Illinois: Scott, Foresman and Co.

CLARKE, A. A. et al. (1977) "The project on teaching undergraduate sociology: core concepts." Mimeographed.

CLEAR, V. (1966) "A new approach to social problems." Improving College and University Teaching 14 (Spring): 93-94.

CLIFTON, A. K. (1976) "Doin soc: an evaluation of actively involving students in the work of sociology." Teaching Sociology 3 (January): 138-147.

CONOVER, P. W. (1974) "The experimental teaching of basic social concepts: an improvisational approach." Teaching Sociology 2 (October): 27-42.

CRAIG, W. W. et al. (1976) "Creative learning through the application of sociological principles." Teaching Sociology 3 (January): 107-130.

D'ANTONIO, W. V. and H. F. LENA (1976) "Preparing graduate students to teach: an ongoing experiment in a small-class setting." Teaching Sociology 3 (April): 277-288.

DUBIN, R. /ed./ (1976) Handbook of Work, Organization and Society. Chicago: Rand McNally and Co.

DUKES, R. L. (1975) "Teaching introductory sociology: the modular approach." Teaching Sociology 2 (April): 165-176.

ELDER, C. D. (1973) "Problems in the structure and use of educational simulation." Sociology of Education 46 (Summer): 335-354.

FULLER, R. C. (1938-39) "The problem of teaching social problems." American Journal of Sociology 44: 415-435.

GAMSON, W. A. (1971) "SIMSOC: establishing order in a simulated society." Simulation and Games 2 (September): 287-308.

GLADDEN, J. W. (1953) "Research oriented teaching of the course in the family." Social Forces 32 (October): 56-60.

GLASER, D. /ed./ (1974) Handbook of Criminology. Chicago: Rand McNally and Co.

GOLDSMID, C. A. (1976) "Resources." Teaching Sociology 3 (April): 339-378.

--- (1977) "The teaching of sociology, 1975-1976: annotated bibliography and review." Teaching Sociology 4 (April): 341-368.

--- (1978) "The teaching of sociology, 1977: annotated bibliography and review." Teaching Sociology 5 (April): 325-360.

--- (1979) "The teaching of sociology, 1979: review and annotated bibliography." Teaching Sociology 6 (April): 291-328.

GOSLIN, D. A. /ed./ (1969) Handbook of Socialization Theory and Research. Chicago: Rand McNally and Co.

GREENBLAT, C. S. (1971) "Simulation, games and the sociologist." The American Sociologist 6 (May): 161-164.

--- (1973) "Teaching with simulation games: a review of claims and evidence." Teaching Sociology 1 (October): 62-83.

GREENBLAT, C. S. and R. D. DUKE (1975) Gaming-Simulation: Rationale, Design, and Applications. New York: Halstead Press/John Wiley and Co.

HARE, A. P. (1976) Handbook of Small Group Research (2nd ed.). New York: Free Press.

HIMES, J. S. (1952) "The factor of social mobility in teaching marriage courses in Negro colleges." Social Forces 30 (May): 439-443.

INBAR, M. and C. S. STOLL (1972) Simulation and Gaming in the Social Sciences. New York: Free Press.

JOURNAL FOR PERSONALIZED INSTRUCTION. Georgetown University, Washington, D.C.: Center for Personalized Instruction.

KOLACK, S. (1975) "A course in ethnic studies." Teaching Sociology 3 (October): 60-73.

KUHN, T. (1970) The Structure of Scientific Revolution (2nd ed.). Chicago: University of Chicago Press.

KULIK, J. (1975) Memo to the Faculty. University of Michigan: Center for Research on Learning and Teaching. No. 55.

MARCH, J. G. [ed.] (1964) Handbook of Organizations. Chicago: Rand McNally and Co.

MCKEACHIE, W. J. (1978) Teaching Tips (7th ed.). Lexington, Mass.: D. C. Heath Co.

NEW ENGLAND SOCIOLOGIST (1978) "Deviance, an undergraduate course outline." 1 (Fall): 40-44.

POOL, I., et al. [eds.] (1973) Handbook of Communications. Chicago: Rand McNally and Co.

PSI IN NOVEL ENVIRONMENTS, 4th National Conference, April 15-17, 1977. Book of Abstracts on the Conference.

RASER, J. R. (1969) Simulation and Society: An Exploration of Scientific Gaming. Boston: Allyn and Bacon.

RILEY, J. W. (1940) "A note on the teaching of social change." Journal of Educational Sociology 14 (November): 139-163.

RUSHKIN, R. S. [ed.] (1976) An Evaluative Review of the Personalized System of Instruction. Washington, D.C.: Center for Personalized Instruction.

SMIGEL, E. O. [ed.] (1971) Handbook on the Study of Social Problems. Chicago: Rand McNally and Co.

SNYDER, E. F. and E. SPREITZER (1974) "State of the field: sociology of sport: an overview." Sociological Quarterly 15 (Autumn): 467-487.

TAVEGGIA, T. "PSI: A Summary of Comparative Research, 1967-1974."

VAN VALEY, T. (1976) "The 'piggyback' structure." Teaching Sociology 3 (January): 160-166.

WASHBURN, P. C. (1974) "On teaching political sociology." Journal of Political and Military Sociology 2 (Spring): 21-32.

WILSON, E. K. (1976) "The carolina course to launch sociology instructors: three features and some general reflections." Teaching Sociology 3 (April): 249-264.

Introductory Comments to
Chapter 8, 9, 10

Everett K. Wilson

These chapters are linked with a common concern: how improve the linkage of ends and means of sociology instruction by recognizing and exploiting differences in student population?

Wilson starts with the assumption that the success of a course depends in part on the people recruited for it. Selection improves the product. Desired outcomes may also be more surely achieved by an adaptation of group ends to the traits introduced by the newcomer. Selectivity in accordance with group (course) goals can be accomplished in various ways: an early warning system through expanded course outlines, syllabi, pre-tests or entrance tests, exposure to final examination on first day of class, division of the group in accordance with interests (with an adaptation of group goals to those interests), and as a last resort, withdrawal of those for whom the group (course) aims are inappropriate.

Heirich expands this theme by suggesting ways of systematically spotting students' interests and expectations in such a way that these student attributes can shape the design of the course. A planned or discovered heterogeneity can be exploited in the service of enhanced motivation and higher levels of learning. Just as the analysis of customer attributes reveals complementarity that can strengthen the course experience, so it can disclose breaks in experience that need bridging, along with the varied resources (student experiences) to help in that bridging. Analysis of student categories leads to the problem of appropriate permutations of program for differing clientele, both in a given course and across a curriculum.

Solomon fixes on the problem: How is learning affected by the interaction of the characteristics of students and those of the situation in which learning is presumed to take place--i.e., on trait-treatment interaction (or what has been called aptitude-treatment interaction (ATI)? Most of the ATI studies have dealt with a few such person-variables as ability and motivation, and such treatment variables as degree of structure in treatment (teaching), degree of supportiveness, and the like. As an example, positive outcomes link higher ability with lower structure. Solomon advances the cause by identifying a range of variables likely to be at work in producing student-treatment interactions: environmental variables (institutional, classroom and teaching styles); student attributes, cognitive and motivational; student responses, perceptual and behavioral, in class and out; and outcomes as seen in mastery of factual content, comprehension and changes in interest.

Thus, having discussed goals in the first five chapters, and the means and resources available for achieving them (in the just preceding chapter), we turn now to the way differing student populations affect that means-end linkage.

Selective Recruitment for Successful Sociology Instruction

Everett K. Wilson
University of North Carolina at Chapel Hill

> "The outcome of a course, in personal and collective achievement, is largely determined before the first class meets."

We take this proposition as axiomatic: The relative immortality of institutionalized human groups depends on the careful induction of its highly mortal members. But the persistence of human groups entails achievement of their goals. (Indeed, the boundaries that define a group include a specification of its ends.) The goals may be modified over time: the annual March of Dimes against polio is shifted to fight another disease. But a defeated Wehrmacht disappears. A business that fails to sell its product goes bankrupt. A school no longer graduating students can't survive. If, then, its persistence depends on the continuous achievement of a declared mission, recruitment must be geared to the selection of those whose traits promote achievement of group goals. This link between organizational goals and selective recruitment is tersely put in common sayings; you can't make silk purses from sows' ears; garbage in, garbage out. The quality of the product depends on the quality of the raw materials.

As the obituaries in the daily newspaper remind us, human communities are incessantly in flux, hence the imperative of institutionalized recruitment. Family, school, church and government shape the newcomer to adult status. Those who are egregious misfits we ostracize--in jails, hospitals, sanitaria, rest homes. Beyond the surface features of social selection and rejection there are the deep, slow flowing currents of biological evolution, selecting and rejecting over aeons. Biological, like social survival depends on the emergence of selected traits.

I. Selective Recruitment: its ubiquity as a condition of achieving organizational goals

Selection for successful membership in human groups occurs at many levels. As examples, cursorily noted, consider selective recruitment at the national level and in various institutional spheres--the family, polity, religion, and education.

1. At the national level: selection of immigrants

After the post World War I flood of immigrants we were very selective in admitting immigrants to the United States. For they would perform effectively as citizens only if their purposes, and the skills for their realization, resembled those of natives. And so our Congressmen agreed that we should control immigration so that imported social and cultural characteristics could be readily shaped to match those of the native population. /Although it raises tangential problems, we should note the assumption that what is, is good. Or what's familiar is better than what's strange. Or what's most frequent, what constitutes a norm, statistically, is what is normative, evaluatively./

We can make a rough check on the extent to which, under our National Origins scheme, recruitment of immigrants was so managed as to match the traits of newcomers with those of oldtimers. We can find 20 countries (nationalities) on which we have both Bogardus social distance scores and the percent the country's quota was of all immigrants admissible.[1] Estimating the relationships between the rank ordered distributions we get a Spearman's rho of .83. This certainly suggests that selective recruitment is very much in accord with the stipulations of the organization's members.

2. Family: selection of children on the basis of sex

The family is a prime example of recruiting newcomers to fit organizational preference. Consider a family as an organization whose members have hopes and purposes that are affected, depending on the number and sorts of persons it recruits. It is an organization that transforms today's child into tomorrow's adult, one who bears the appropriate freight of values, skills, beliefs.

At the point of admission, we think of recruits' attributes as being determined only by one criterion, number. Parents may choose whether to admit no other members, or one or two--or more. Yet even as to traits we'd think impervious to selective influence, we operate to meet organizational standards. One such trait is the newcomer's sex. Unlikely as it seems, the Freedmans and Whelpton (1960) have alerted us to sex as a criterion for seeking new recruits. They find reason to suppose that, in our society, parents prefer having children of both sexes. (With children of both sexes, parents have a chance to recapitulate and savor their own sex-linked development. Since so many activities are sex-segregated, having children of both sexes broadens parents' acquaintances and associations. And for the shortcomings of his own experience, a parent may look to his same-sex child for compensatory achievements.) And so it follows that parents having children of one sex, alone, will be more inclined than others to have an additional child--and to stop when the child of the missing sex is added.

The data bear them out. And their findings held when duration of marriage, wife's religion, education, and the extent of her work experience were held constant. In no case was the direction of the relationship altered and in only one case, religion, was the extent of the relationship affected. Thus in a modest way (and in the aggregate) the family may alter the balance of sexes in its organization through the recruitment process--deciding whether to have an additional child.

3. The polity (military): ascriptive to achievement criteria

Perhaps organizational selection is most rigorous when the group product touches the public interest. In wartime, ascriptive criteria go by the board; for to follow peacetime patterns of recruitment would compromise the product, destruction of the enemy. Never had so many women entered the labor force as in World War II. In unprecedented numbers, women and blacks served in the military. A societal crisis altered recruitment patterns in accord with compelling organizational (societal) goals. (There is probably a finer selective process at work between branches and divisions of the various services. One might hazard the guess that age, education, and motivation will vary depending on whether one is recruited for early or late wave, close or distant killing.) Self- and organizational selection affect both personal and organizational achievement.

4. Religion: the importance of selective recruitment for entering a religious vocation

Entering some groups is societally, organizationally, and personally a costly business. Medicine, law, teaching--and in this example, the religious --are cases in point.

Committing one's self to such organizations is costly not only because the training is long, but also because of aborted entry. For example attrition rates may be very high, depending on the effectiveness of self-selection and organizational assessment. Coville reports (1968) that at St. Vincent's Hospital School of Nursing, there were 5,039 applicants over a 12 year period. Of these, 2,200 were admitted and about 90% graduated.

But the average attrition rate for seminaries and religious communities ranges "between 40 and 50 percent, and in some instances extends to 75 percent" (Coville, 1968: 37).

To prevent such losses, costly for all concerned, the Committee for the Study of Methods in the Psychological Assessment of Candidates for the Religious Life has identified aspects of the religious vocation that require careful evaluation of recruits. These, according to Coville (1968: 14, 15) are: a spirituality that leads a person to "concentrate primarily upon the core duties of the religious vocation, rather than on its devotional aspects, social advantages, or even opportunities for secular achievement . . . appropriate motivation, adequate intelligence, relative emotional stability, a capacity for effective communication, a good interpersonal relations, and sound physical health."

Of course it is not only the attrition rate that is at issue, although that is serious enough. Presumably effective recruitment can also reduce the incidence of weak and ineffective performances in role.

5. Education

This chapter leads to the final suggestion that success of a sociology course depends in no small measure on student traits brought to, or specified as requisities for it. But organizational aims and selective recruitment operate at other levels--colleges and universities, and departments--as well.

Karabel and Astin (1975) find that both SES and academic ability operate to select and sort students into colleges of differing quality (measured by mean academic aptitude of entering freshmen and revenue available per student). Academic ability appears to be more influential than SES, although the latter exerts some independent effect on the sorting process. The authors write (Karabel and Astin, 1975: 395):

> What seems to have developed in American higher education is a differentiated system of colleges and universities which bear a close resemblance to the tracking system of American high schools. . . . Comparing institutions of the highest and lowest levels of ability, we find that differences between levels far exceed ability differences within levels. . . . and that students in the lowest track--two-year community colleges--receive an education which is often different from and shorter than what their counterparts in four-year colleges receive. Meanwhile, students at elite colleges continue to be channeled into the professional and graduate schools which

> give them entree into many of the choicest positions in the occupational structure. This implicit tracking system within higher education, though little noticed, is becoming every bit as real as its counterpart in the secondary school system.

This reminds us that the differentiation of organizational mission and selective recruitment may not be matters of conscious planning but outcomes of societal structure--in this case, one in which SES is strongly related to privileged preparation and modestly related to SAT scores. It reminds us, too, that the issue of personal and organizational success as a function of selective recruitment is a matter quite different from a judgment on the worth of the mission.

Sociologists hardly need to be reminded of the selective influence of family background on school achievement. At the secondary level these forces are at work, manifesting themselves both in curricular tracking and in attrition rates. The influence of background factors appears to be greater than differences in school organizations, themselves. Coleman writes (1971: 24) that

> Level of skills attained through . . . schooling differs for blacks and whites--in part because of the different school resources to which blacks and whites are exposed, but even more because of the different family and community resources that support or constrain the child's achievement in school /emphasis added/.

Coleman makes an estimate of the relative effect of family background factors and organizational characteristics (those of the student body, teachers, and school facilities) on measures of verbal achievement of 12th grade students. Findings are given in Table 1. It should be emphasized that family factors are probably underestimated here.[2] Even so, we see that what goes into the school from family background is more powerful than any one of the three in-school resources. The input variables represented by the family have almost twice the effect on output (verbal achievement) that teachers have, or school facilities; and about the same effect as measured variations in the student body. (But one should note that fellow students are the instruments through whom similar background features are brought into play.)

Table 1[a]

STANDARDIZED REGRESSION COEFFICIENTS, SHOWING RELATIVE EFFECT OF VARIATIONS IN FAMILY, STUDENT BODY, AND SCHOOL ON VERBAL ACHIEVEMENT OF 12th GRADE STUDENTS

Resources	Negroes North	Negroes South	Whites North	Whites South
Family	.23	.22	.34	.34
Student body	.21	.23	.09	.11
Teachers	.13	.12	.09	.04
School facilities	.13	.07	.10	.07

a. Source: Coleman (1971: 45)

Of course such data simply echo the earlier findings of Coleman et al. (1966) that organization effects (school) are small compared to the powerful impact of the traits brought to the organization. It was in this much debated study that investigators summarized their findings by saying (1966: 325) that

> Schools bring little influence to bear on a child's achievement that is independent of his background and general social context. . . . This very lack of an independent effect means that the inequalities imposed on children by their home, neighborhood, and peer environment are carried along to become the inequalities with which they confront adult life at the end of school.

As a final example from the realm of education, we see selective recruitment at work in the admission of graduate students. Consider the data in Table 2.

Table 2

DIFFERENCES IN GRE SCORES BY STAGES OF THE ADMISSIONS PROCESS, DEPARTMENT OF SOCIOLOGY, UNC-CHAPEL HILL, 1972-77[a]

GRE Category	Mean difference in scores, 1972-77, between those Applying and Admitted	Admitted and Enrolled	Applying and Enrolled
	(I)	(II)	(III)
Verbal	+60	-11	+49
Mathematical	+60	-28	+32

a. Data from the Chairman's annual report.

Following the Roman numerals we note:

I. Standards for admission were so set as to eliminate poorer scorers. Over these six years there is an average gap of 60 points between mean scores of applicants and mean scores of those admitted.

II. This organization does not stand alone; it operates in a competitive field of organizations. There was selective attrition among those admitted, so dropping the average score of those enrolling by 11 points on the Verbal GRE and 28 points on the GRE in Mathematics.

III. Net selectivity, then, is reduced. Difference in mean scores between those actually enrolling and those who applied, over these six years, is +49 for the Verbal and +32 for the Mathematical GRE.

Thus we see in this instance both personal and organizational selectivity at work. The organization stipulates needed skills, the ability to manipulate symbols, verbal and mathematical, for admission to a sociology department. Also, but in competition with like organizations, it stipulates the level of desired performance. The person has first selected the field of concentration and then the organization to which he will commit himself for training. (Other elements, obviously, enter into the union of person and group.)

6. Sports

Rivalry, competition, conflict seem to engender refined and explicit criteria of selection in the service of personal and organizational achievement. This is especially the case with professional and preprofessional sports: criteria for recruitment are meticulously specified and measured. Much is invested in selective admission of athletes to teams. Younger athletes, those in feeder schools, are identified and scouted. Dimensions of performance are operationally defined and measured in infinite detail. Organizational needs are specified. One measure of the care taken in matching newcomers' talents to organizational needs is the dollar investment in such personnel work. For example, among the costs of the UNC-CH football program in 1977 (1.2 million) there is an item of $85,000 (7%) set aside for recruiting. It is quite clear that organizational success hinges on traits and talent brought to the organization through careful, competitive, selective recruitment. The outcome, the record of wins and losses inscribed in the fall, is in large measure the result of activities that occurred long before the person joined the organization.

Table 3 summarizes this short roster of examples of ways selective recruitment serves organizational ends. The account is cursory, the examples abbreviated. Self-selection and organizational selection are not clearly distinguished, much less weighted as to their effect on organizational purpose. Nor have I clearly discriminated between calculated mechanisms of selection and those deeper currents in social life that fix our range of options. Even so, it seems clear that social organizations, having permeable boundaries, typically rely for their preservation on selective recruitment. We can add that recruitment will be effective to the extent that it promotes the achievement of group purpose.

Table 3

A SUMMARY OF INSTANCES OF RECRUITMENT IN SUPPORT OF ORGANIZATIONAL GOALS

Type of Organization	Goal or Purpose	Mechanism of Selective Recruitment
1. Nation	resemblance to natives in commitment to capitalism, Judeo-Christian tradition and the like	Immigration quotas
2. Family	children of both sexes	having additional child and stopping with advent of other-sex child
3. Polity	destruction of enemy	testing/allocating mechanisms including adtecedent education
4. Religion (religious vocations)	strengthening the church, commitment to service of communicants and others	psychological assessment
5. Education (college & university)	differential training for a differentiated labor force through schools differing in quality	scholastic aptitude, costs
Education (secondary)	preparation for work, for higher education, for common citizenship	family and other background differentials, tracking
Education (graduate Dept. of Sociology)	training in skills of sociological inquiry	GRE scores
6. Sports	winning teams	observation & records of performance in feeder schools

It will not have escaped the reader that there is another way of matching organizational mission with newcomers' traits. (I have emphasized only the selection, or input of appropriate traits.) But another means of achieving compatibility is to alter the purpose and requirements of the organization. To illustrate the two extremes, in a classroom situation, imagine first, the organizationally nonadaptive position: a course in stratification admits only junior level students with a background in statistics and who are able to separate structural effects (shifts in the labor force) from those of familial status and parental occupation. At the other extreme there is the course in stratification without prerequisites in which each student is encouraged to pursue his own interests even though they relate most tenuously to class structure, its sources and consequences.

This suggests that there is a certain point of adaptation to the attributes of recruits beyond which an organization cannot go without sacrifice of its purposes--or even its integrity. The bounding membranes that define a group, its distinctive ends and means, are always permeable. Retirement, death, resignation, mobility and replacement through birth, migration, and promotion are normal osmotic processes. Inevitably newcomers bring novel infusions that work modest permutations on organizational ends and means. But if the cell's walls are ruptured, if it is not a matter of permutations but of transformation, then the cell dies. To spell out the analogy, the old organization disbands or becomes a new one.

At the organizational level, this was precisely what happened to some schools during the late '60s and early '70s. Recall the vacillating course followed by the captains of what was then a growth industry.[3] The tendency was to shape organizational purpose to individual inclinations of the recruits. How often did it happen that, where student feet crossed campus life, the blacktop was obligingly laid down? (And if the organization lacked clearcut purpose, how many paths must crisscross the 4 years of a student cohort's experience?) Respect for the person, the individual as a final arbiter--these pieties found ideological support in market place and ballot box. But when, with virtuous intent, individual inclinations supplanted common purpose, the organization as an identifiable entity disappeared.

The same may be said of the smaller and shorter-lived organization that is a course in sociology. Indeed, all this is prelude to distilling a few propositions that might profitably be applied to our teaching of sociology. What are the propositions? And what do they imply for our instruction?

II. Propositions Implied by the Counterpoint of Personal and Organizational Goals

I've offered a grab bag of examples. In biological and social evolution we see selectivity at work, linked to personal adaptation and corporate survival. Self selectivity and social selectivity operate in recruitment for national membership, entry into the professions, including sociology, recruitment for military service, for sports--even for membership in the family. What can we make of such instances?

It would be useful if we could deduce a set of propositions which might, then, have some implications for teaching sociology. My first draft of propositions follows:

1. Definitional: An organization is a distinctive structure of relationships in the service of an end.

2. That it is distinctive implies a plurality of organizations from which it differs.

3. Most, if not all organizational boundaries are permeable: deaths, resignations, promotions, graduations, require replacements.

4. Any measure of organizational stability is some function of kind-for-kind replacement: entrance matches exit. (In this and the succeeding propositions I'm setting aside extra-organizational influences on stability or instability.)

5. Any measure of organizational change is some function of a mismatch between traits of those departing and those of new recruits. (I use the word, mismatch, in a descriptive, not an evaluative sense.)

6. Organizational change always and inevitably occurs since, at a minimum, each recruit's attributes emerge from a unique social ancestry.

7. Beyond the age of compulsory schooling, recruitment occurs in a matrix of competing organizations. (Range of options and extent of competition is probably a curvilinear function of age, being least in childhood and in the later years.)

8. The organizational environment changes and, in doing so, alters a given organization's mission. Thus the operations of a given organization are qualified both by its exit-entry balance of traits and by extra-organizational influences.

9. Recruited traits which are a random selection from the available population (in contrast to a calculated orchestration of traits) must be eccentric to organizational needs and will work against the achievement of group goals and those of some individuals (as is true of some outstanding students in open admission situations).

10. Recruitment to ensure achievement of group goals will entail:

    i. a precise specification of group goals, and their connection with individual goals,

    ii. clear communication of those ends--and the means needed to achieve them,

    iii. some reasonable adaptation of organizational means in the light of the changing attributes of recruits,

    iv. some adaptation to extra-organizational demands so as to exploit the relevant outcomes of other groups,

    v. inevitable error in recruitment, so requiring the safety net of withdrawal and reallocation.

III. Application of such propositions to the classroom group

The obvious point is that sociologists' knowledge of social organization might have some implications for the working group committed to a course. As with any other organization, achievement in a sociology course depends partly on a reasonable compatibility between the intentions of the leader and those of group members. It hinges, too, on some common understanding of the means employed to achieve the ends. And as a last resort it depends on a means of withdrawal of those who tardily discover that they have drifted into a group whose mission is at odds with theirs.

1. ENDS: specifying organizational (course) goals

There are informal influences on recruitment: parent or alumnus recommends taking something in the social sciences, or a course with Professor X; and campus scuttlebutt doubtless plays a part in steering prospects to--and from--particular courses. But as sociology instructors we would seek systematic procedures that increase the chances of student achievement and course success--which is to say, achievement of instructor. Hence we will think about the common, and uncommon, ways of recruiting a crew with compatible ends in view. Consider, for example, the catalogue statement; or beyond that, more detailed course descriptions; a readily accessible file of syllabi in the department office; recent student evaluations of the course; an available file of the last three years' final examinations; or a file of entrance examinations or pre-tests for the course; or, more traditionally, permission of instructor. A final, and often too late, specification of course ends lies in the syllabus handed out on the first day of class; or the introductory statement of the instructor.

There are two means, conventionally used, of revealing the nature of a course to prospective members, the catalogue statement and the syllabus. The first is likely to be cryptic and therefore inadequate. The second is likely to be too late (unless there is available to the prospective student a department file of course syllabi); and in any case is seldom written in view of this special problem: it is addressed to someone who, by definition is a novice and doesn't know the score. The reasonable inference, then, is that the initial (recruiting) statement should avoid technical abstractions; it should be clear and concrete. It will not do to prate pious vagaries such as this: "An aim of this course is to convey a sense of the sociological perspective." (The last two terms have no intelligible referents in the newcomer's experience.) Nor will recruitment be effective if we rely on the terse, abstract and uninformative blurbs that befog college catalogues. On the contrary, a statement of organizational goals should specify precisely the skills, knowledge and attitudes which the course aims to achieve; and to do so in a way that accurately samples the range of work spanning a semester's time. Neither effective self-selection nor organizational selection can come into play when group goals are put in the instructor's professional lingo.

To the extent that we fail to communicate clearly to our likely recruits, course goals and the means for achieving them, we risk cross purposes. Incompatible ends and means put the course's aims in jeopardy, lead to poor instructor evaluations, lower the department's reputation on campus and, to some modest extent, demean the discipline.

Clear communication of group ends (and means) suggests that we might do what we would with a research instrument--give it a dry run. We would find at times that abstract jargon strains the reader's wits--and perhaps, as Meadian theory would suggest, dulls the writer's! (We should remember the words of Richard Mitchell who edits the Underground Grammarian at Glassboro State College: "Bad English is a form of malpractice in an academic professional." Perhaps the time will come when malpractice insurance will protect us against the adverse outcome of recruiting to an organization newcomers whose ends are at odds with those of the group--because they could not understand catalogue description or course syllabus.

Of course the newcomer to an organization is always at some disadvantage. Even good will and high skill are inadequate to convey, fully, the pains and profits of the impending experience. Under these circumstances it is critically necessary that the student understand not only the goals of the course but their justification. As we would think it inappropriate to impose our tastes in food and drink on others, so would we be reluctant to impose an intellectual diet without reason. Hence our obligation to say, for example, that since the self is socially shaped, a student can come to understand himself better as, through sociology, he explores the social sources of his selfhood. Or, again, since our knowledge is only as sure as the methods we use to get it, the attention sociologists give to methods can strengthen his assurance as to what he knows--and doesn't!

The academic diet sometimes gets less consideration than the selection of food at a cafeteria counter. If recruitment is to be successful, the instructor (personnel manager) needs to set in motion the process of self-selection. How is a student to know whether a course in sociology is appropriate for him?

One answer, which I'll return to later, is that a course is appropriate if the student has met stipulated prerequisities; for there is, presumably, some sensible link between a given course and other(s) specified as its prerequisite.

Another answer is to provide more detailed information than is available in the catalogue description--or in the syllabus which he may not see until the first day of class (and sometimes later than that, if the instructor has been delinquent). At UNC-Chapel Hill the Department makes available in the central office an expanded description of each undergraduate course offered. The extent to which students know about or use these expanded course descriptions is uncertain. A second more elaborate version of this device is found in the Political Science Department at Oberlin. A faculty-student committee provides a mimeographed brochure for each course offered in the Department. For each course there is a description of its aims and procedures, a summary of student evaluations of the course and a statement by the instructor. One might suppose that this must promote intelligent self-selection.

Another way of informing the student about the purpose of a course is to provide him, ahead of time, with a copy of the final examination. Learning, one assumes, is the general aim; and that means that a student can do (answer) at a later time what he could not do earlier. Desired performance at that later time is spelled out, operationally, in a final examination. In so specifying course goals we would be following another injunction that informs our work as sociologists, the translation from abstract theoretical propositions to the operationally specified indicators necessary for testing. The test of a physician is in the ability to speed healing through artful diagnosis and prescription. The test of the football player is his ability to gain ground against his opponent. Such aims can be quite clearly put and readily understood; for they are expressed as resultant behavior. We might be equally effective in defining goals were we to exploit final examinations to this end.

Perhaps the most obvious recruiting filter is an entrance examination. (One might make the case that entrance examinations certify a readiness and legitimate an implicit contract as no other device can do.) Not only do they define course aims in operational detail. They provide a benchmark for ultimate assessment of course gains. And they reveal strengths and deficiencies to guide student and instructor in shaping subsequent course work. Immediately administered in a drop-add period, they help both parties determine whether the newcomer should continue; or substitute some more appropriate course experience.

In less determinate fashion, a course syllabus given students on the first day of class may serve a similar purpose. Its effectiveness as a selective mechanism depends, again, on the clarity and specificity of its statement.

2. MEANS to be employed in achieving goals

The procedures to be used, no less than the goals they serve, may offer useful, selective clues. Recruitment will serve personal and group goals better to the extent that student and instructor can answer the questions: Is the newcomer adequately prepared? Will the instructor shape course procedures to exploit familiar experience on which I can build?

May I employ other, concurrent work to advance my understanding in this sociology course (and vice versa) so lending a degree of integrity to my education?

A typical hurdle for admission to a course is the requirement of an antecedent course. To achieve both personal and group (course) goals, the instructor announces: you must come equipped with--what? The typical answer is cryptic. We specify another course (often, simply by number) as prerequisite. Too often communication is impeded as a murky description of one course specifies a preceding, murkily prescribed course as prerequisite. Things are further complicated when the preparatory course is taught in many sections, by different instructors, using different materials. Under such circumstances, the preparation of newcomers will vary widely, and in unknown ways. Prerequisites can serve selective recruitment effectively only when desired outcomes from the antecedent course are put as demonstrable abilities.

The fit between person and group may be achieved in two ways (or both): by recruiting students who have the means to handle course procedures, or by adapting organizational means to recruits' characteristics. Adapting organizational means to the experience of the recruits is critical when schools vary as do ours in the United States. It is not the same thing to teach sociology at CUNY, Queens as at Vanderbilt. The New School for Social Research is a far cry from Hopkins, as both are from Reed, Oberlin, or Grinnell. And all differ in various ways from the Junior or Community College, or from Tulane or Antioch. There may be no need to modify organizational goals in accordance with such marked differences in membership (student populations). But some reasonable adaptation of means is indicated.

One might adapt the means of instruction by tapping students themselves, both for theirs and their families' characteristics; and for dependent variables represented in opinions and attitudes. (See, for example, the first stage in Paul Baker's "Social Awareness Project," Baker, 1975.) Such data, collected early, can be periodically used both descriptively (and for comparisons with state, regional, and national data); or as creating a point of inquiry: What explains the difference between these attributes and those of other populations? One might contend that some such tapping of resources represented in student experience is imperative; that the means used to reach organizational goals must somehow touch the experience of the new recruit. Otherwise, to invoke Mills, biography and history never intersect.

One speaks into the wind. One synapse doesn't know what the other is up to. Relevance is not for titillation, of course; it is for communication, in the basic sense of a community of experience enabling propositions to span a wider range of empirical reality. Such shaping of means to the peculiarities of group membership may be critical where the pool of recruits differs markedly from those elsewhere. (It may be critical, too, as the group leader ages and successive recruits differ markedly from those he first knew.)

There is a means of matching group goals and newcomers' needs that combines the two processes so far mentioned (selective recruitment and organizational adaptation to newcomers' traits). It is the creation of two or more class groups on the basis of traits revealed upon admission. Thus, for example, faced with a large, heterogeneous aggregation[4] of students, Paul Baker responded by creating three distinct organizations, each with somewhat differing ends and means which then became the criteria for student self-selection. (See Baker et al., 1974).

Organizational success, the success of the course, depends not only on self- and group-selection; nor alone on organizational adaptation to the peculiarities of a population, or some combination of these. It depends, also, on extraorganizational influences. Organizations exist in a competitive field. (This was illustrated in the drop in GRE scores between graduate students admitted and those ultimately enrolled.) And a course operates in a set of courses whose influences intersect (or fail to intersect) in our sociology classrooms. (Clearly it is not other course memberships, alone, that do--or might--affect the outcomes of sociology instruction. The influence of family and faith and employment intersect here, too. But here I'm taking concurrent course experience as an example.) Adaptation to extraorganizational demands suggests the possibility of exploiting the outcomes of other group memberships.

Thus another prescription as we adapt organizational means to the traits of recruits. We need to know something about other academic experiences shared, concurrently, by a number of persons in our group. This is not simply a matter of enhancing the integrity of a student's learning experience. It is also a matter of extending the meaning of sociological inquiry for people most of whom will reap the gains of our course only in connection with some other activity. There are points at which Shakespeare and algebra, biology and logic touch sociology. And the other way around. Goals of the group (course) and of its members are more readily achieved when relevant aspects of the organizational environment are identified and made to serve instruction in sociology.

3. WITHDRAWAL (as safety net for recruitment failures)

Finally, since any group will sometimes err in admitting and some will inevitably err in entering the group (some registration procedures more resemble Russian roulette than rational shaping of educational experience) there should be the safety net of withdrawal. A first week pretest can serve both as a base point for measuring change and as a means of letting the student know more surely where he is and whether he should be there. A review of the final examination will serve the same end. Early review of typical reading material will help the student make a decision. A set of handouts or laboratory exercises can foreshadow future demands. And so, above all, will a carefully wrought syllabus that speaks clearly--not in tongues--to the novice.

In sum, we can better sociology instruction by acting on this knowledge: organizational outcomes are crucially affected by selective recruitment. So, likewise, with a course. Its outcome, in personal and collective achievement, is largely determined before the first class meets. Much depends on the mix of talent, will, and motivation that comes from the registration line to the first course meeting.

But we need not accept the fact of registration as the fate of the course. Human groups are typically, if not invariably, selective. In our sociology courses we can put selectivity to work to yield a certain minimum preparation for what is to come; together with a certain optimum heterogeneity of experience that can be exploited in the analysis of differing social patterns. And so, using our inferences from organizational analysis, we can promote selectivity in various ways.

We can offer a clear, precise specification of course goals--in syllabi and expanded course descriptions--instruments that should be pretested and, after revision, made easily available to students. We can also make available recent student evaluations of the course (if thoughtfully done), and final examinations as operational statements of course goals. We should include carefully crafted statements justifying those goals.

Similarly as to means, the skills and knowledge need to get on in the course. To this end we can provide behavioral statements of prerequisites (not simply another course number). Without compromising course goals we can adapt course means, employing data gathered from and on group members, occasionally using ideas drawn from their other courses, concurrently taken, and gearing ideas and data to the occupational categories that reflect students' expectations.

Helping students discover degree of compatibility between their traits and course goals may provide sensible reasons for withdrawal. Or the same instrument, such as a pretest or background questionnaire, may provide the instructor with data needed to fit differing means (course experiences) to different categories of students. This, in effect, is a modified withdrawal, shutting off inappropriate experiences while exposing the student to those that promise a higher yield.

All of this reflects a conviction that a careful analysis of organizational recruitment can help us identify aspects of classroom organization which, if attended to, would improve our work. And since the analysis of social organization is our métier, it is to say, in short: physician heal thyself.

Notes

1. I draw here on data developed for a student exercise in Wilson (1971: 13-15).

2. There are three reasons for this, two of which Coleman mentions. First, student responses about family background probably vary less and are more skewed to the favorable than were the other data on the school situation provided by teachers and principals. Second, blacks were compared with blacks, whites with whites, reducing background measures of variance. Finally, the influence brought to bear from the student body (listed under what Coleman calls "Resources") are themselves a reflection of the same background that helps or hinders school achievement.

3. Adapted from Wilson (1968).

4. It is plausible, of course, that heterogeneity might be a positive criterion in recruitment. This is a special emphasis in the following chapter by Max Heirich. A significant line of research would pursue the question: Under what conditions are heterogeneity (and homogeneity) in specified student traits advantageous in achieving what learnings?

References

BAKER, P., BAKSHIS, R., and W. TOLONE (1974) "Diversifying learning opportunities: a response to the problems of mass education." Research in Higher Education, Vol. II. APS Publications, Inc.

BAKER, P. (1975) "Social awareness project." Teaching Sociology 3: 74-80.

COLEMAN, J. (1971) Resources for Social Change: Race in the United States. New York: Wiley-Interscience.

COLEMAN, J., CAMPBELL, E., HOBSON, C., MCPARTLAND, J., MOOD, A., WEINFELD, F., and R. YORK (1966) Equality of Educational Opportunity. Washington, D.C.: USGPO.

COVILLE, W. (1968) "Basic issues in the development and administration of a psychological assessment program for the religious life," in Assessment of Candidates for the Religious Life. Washington, D.C.: Center for Applied Research in the Apostolate.

FREEDMAN, R., FREEDMAN, D., and P. WHELPTON (1960) "Size of family and preference for children of each sex." American Journal of Sociology 66: 2.

KARABEL, J. and A. ASTIN (1975) "Social class, academic ability, and college quality." Social Forces 53: 3.

WILSON, E. K. (1968) "Disappearing schools and scholars." The Antioch Review 28: 2.

--- (1971) Instructor's Supplement to Sociology: Rules, Roles and Relationships. Homewood, Illinois: The Dorsey Press.

Assessing the People We Teach: Aids to Course Planning[1]

Max Heirich
University of Michigan

With the right students, those with pertinent interests, skills and experience, classes sometimes come alive. Those who have seen a solid nucleus of students get excited about a course, the excitement then spreading to others so that the daily encounters and the work done become deeply engaging, know how satisfying teaching and learning can be. They have also probably discovered what helpful teaching resources the students in a class can be once they come alive to the subject matter.

Many who are fortunate enough to have such a class from time to time find that it happens in a course dealing with topics the teacher finds especially interesting and in courses which students have freely and purposively selected. Rather few of us have any hopes that this kind of learning can emerge within required courses, large enrollment courses, or those not dealing with our own favorite subject matter. We tend to view these unusually engaged classes either as a result of an effective grapevine (so that the right students find out about a course) or as a random and mysterious happening. If a colleague regularly has engaged classes, we tend to assume that he or she has an innate gift for teaching that cannot be transmitted to others.

In this chapter I will argue that engaged learning of the kind I have just described, although it cannot be guaranteed to emerge at any particular time, can be planned for and not left simply to chance. A few, systematic ways to assess students' interests and expectations can let one build into a course experience sequences of learning, and opportunities for interaction that increase the chance that a course will come alive for an important segment of the class. Moreover, one can do such planning prior to encountering the specific set of students who will be enrolled in the course, by using systematic evidence about the student body and about the particular streams within it that take a course over a period of time. I will suggest a few ways to make this kind of planning assessments and then will describe some general strategies for using such knowledge to structure the content of a course and also the various teaching-learning approaches used, so as to maximize course impact. That discussion will not involve cookbook recipes for planning various courses; instead it will give a few suggestive examples that could start one thinking about the specific course and learning setting that is of immediate interest to the reader.

As I see it, making a potential impact assessment as a part of course planning involves five kinds of activities. First it entails an analysis of the organizational routes through which students are recruited into various sociology courses. This lets one predict common areas of concern, skills, interests, prejudices, or blind spots that are likely to be found within a particular set of students because of the settings from which they come. (If one can find points of connection between the real concerns of this set of people and the real concerns of sociologists, a more fruitful dialogue is likely to result. It also lets one put appropriate information into the recruiting channels so that better matching occurs.) Second, such planning involves an analysis of kinds of exposures and experiences that may be needed so that questions addressed in a course become more than academic topics to students. One can then anticipate types of resources likely to be found within the student group itself for providing such encounters, and mechanisms that will be needed to let these kinds of experiences occur. Third, it means assessing the kinds of knowledge and skill-teaching resources individual students might be expected to bring to the class or to have available to them, so that they can become active participants in the teaching process as well as receptive learners. Fourth, it involves assessing the kinds of diversity likely to be present among students and alternative strategies for dealing with this. Finally, it involves assessing the range of learning situations available to students on a departmental basis.

I. Organizational Routes by Which Students Arrive in Sociology Courses

Selective Recruitment by Schools. Universities and colleges do not draw students randomly from the available college-age population. Karabel and Aston (1975) have documented national patterns that locate students among various kinds of colleges and universities, not simply on the basis of previously demonstrated academic ability, but also in terms of social backgrounds. Everett Wilson (Newcomb and Wilson, 1966) has suggested a series of student attributes that affect how one learns sociology and that are not distributed randomly among undergraduate populations. Medsker and Trent found patterns similar to those described by Karabel, Aston, and Wilson.

An initial assessment regarding students likely to take a particular sociology course, thus, might first involve fairly broad demographic analysis of a schools' student body. Are undergraduates primarily upper middle class suburbanites? First-generation college students from working class, ethnic backgrounds? Small-town elite-family descendants? Black bourgeoisis? What mixes of such backgrounds regularly occur in sociology classes? (If transfer students or graduate students are anticipated, in course planning, from what undergraduate settings do these students come?)

The planning implications of such an initial demographic assessment may be obvious. It lets you predict kinds of experiences, mythologies, goals, and intellectual styles likely to be found within the group taking your class. There will, of course, be many unique individuals, but the social settings from which they come offer special frames of reference for viewing the world; a fair proportion of students can be expected to operate within orienting perspectives appropriate to their earlier social locations. If one is clear about what these are likely to be, it becomes easier to frame an initial question for the course that resonates with interests or concerns of class members and to plan a sequence of topics that can make it easier for them to move step-by-step from understandings they have at the beginning of a course to understandings you hope they will develop as the course goes forward. Additionally, an initial demographic assessment, interpreted as suggested here, can alert one to types of reading materials, field trips or other total learning experiences that may be needed to help the course come alive for different sub-groups of students that can be expected to enroll in the class.

A few examples may make the point clear enough to apply to a wide range of sociology courses. A course on social stratification may be resisted or may be primarily an academic discussion for a group of students who all share the same general social class background unless the course builds in a set of field experiences that make the subject matter real. In contrast a stratification course that includes articulate members of several social classes may become an exciting exploration of social structure and process simply by making sure that the various class members discover one another's backgrounds and perspectives.

A course concerned with the sociology of religion will encounter early problems of communication and comprehension if it regularly includes two sets of students: those hostile to religion or who lack in their own experience anything they would call religious experience; and another set of students who could be called true believers. An initial section of the course dealing with the phenomenology of religious experience and response--or early assignments that take students out into the field among religious communities --could make the questions sociologists ask about religion much more real to both of these sets of students. Their own differences in approaching the topics then could be a rich source of mutual learning. Similarly, a sociology of religion class taught in a denominational college that attracted persons of similar religious experience and outlook--or that attracted a sizeable nucleus of such people--might have to spend some time expanding the experiential base of students before the questions that sociologists ask about religion would make any sense to some of these students.

A research methods course might gain impact were it to explore topics suggested by class members or, alternatively, if it were to pursue a question of known relevance to the concerns of population groups that regularly take the course. Statistics assignments growing out of data and topics with which students have had direct encounters can become engrossing rather than perfunctory learning experiences, if the students are invested in knowing what the answer really is.

The backgrounds and expectable previous experiences of students taking an introductory sociology course will affect the ease with which students relate to questions about the macro-forces at work in social life. These backgrounds can alert the instructor to micro-level questions that should touch the experience of these students in a way that makes social patterns curious or problematic for them. Rightly chosen, such micro-analytic questions could raise further questions in students' minds so that they move progressively, step by step, to focus their attention on larger social forces at work.

Assessment of typical demographic streams that send students to our courses also might alert us to mutual planning and exchanges that could make for beneficial cooperation between various educational institutions in an urban area. For example, a stratification course might work better if jointly taught between an elite or upper middle class institution and a community college.

Internal Streaming. Within a college or university, graduation requirements may stream particular clusters of students into sociology courses. A nursing, education, or engineering program, for example, may require its students to take at least one sociology course, or a particular course title. Similarly, graduate courses frequently have identifiable clusters of students from other degree programs (e.g., public health, psychology).

In my experience, such streams typically are approached by sociologists with dread: it is assumed that these students will have little interest in sociology and little aptitude for it; such students often become either passive or hostile elements within a class. With proper attention, such streaming could become a part of class planning: pre-nurses could be directed into field projects or research literature directly related to their anticipated work in hospitals or with the sick; engineering students could be steered quickly to computer studies, measurement problems, or industrial issues that respond to their passion for precision and for problem-solving. And so forth. If we can once persuade these students that this course will be relevant to their own interests, rather than just a requirement to graduate, they can become a valuable resource to the rest of the class, since they are likely to have skills and experiences which other members of the class lack. Equally important, if they can become engaged in the class, rather than passive resisters, the whole tone of the class will change.

One must be careful, of course, not to make hard and fast assumptions about the motivations relevant to students who come into one's course from various streams. Nursing or engineering students may value a sociology course as a change of pace from their usual subject matter and resent being pushed back into the world from which they have just emerged; an inner-city black student may want, above all else, to examine questions for which race and class are not the primary explanatory variables; or the pre-nursing student may as yet have had no direct experience with health care settings. In planning, therefore, one would not automatically track specialty students into assignments geared especially for people from their backgrounds. Rather one would include these options as possibilities, or would bring such material into the class sessions from time to time.

A look through the college or university catalog, and an examination of computer print-outs for class registration will make clear the forced streams that should be taken into account. In addition, there may be other, more informal streams that are discovered only by student comments, observation, and information directly elicited from the students. In my university, for example, the introductory course gets not only these required streams but a predictable band of athletes, better than the university's proportion of black and third world students, plus a large number of pre-med and pre-law students. This becomes a formidable motivational challenge if one tries to treat the course as a homogenous learning experience, but it can be a rich resource if one plans it with these backgrounds in mind.

Gateway Personnel. Student streaming normally is channeled by a relatively small number of gateway personnel. Some are counselors in special programs (e.g., nursing, engineering) or in the undergraduate counseling office. Others are student influentials in particular population streams. A third, heavily used route, of course, is chain-recruiting through informal student networks spreading from currently enrolled students to their friends. Systematic information about planned courses, their foci, style of teaching, and the like, fed to these information nodes can help improve the fit between students and courses. Assessment planning in this area should involve not only analysis of formal information channels but also the most used informal routes that are easily accessible to instructors or to current students. In some schools departments make available brief, paragraph descriptions of courses to be offered the next term, including such information as who will teach the course, the topics to be addressed, the style in which the course will operate, required work, and the like. Such information, circulated to critical nodes in the formal and informal information networks, can help create a better fit between students and available courses.

II. Analyzing Needed Exposure and Experiences to Make Course Issues Real

There are important distinctions to be made between a touchy-feely course that ignores or waters down hard intellectual issues, one that insults the intelligence of students by assuming they cannot recognize the importance of its subject matter when they already are prepared to proceed with the analytic task at hand, and a course that genuinely needs certain kinds of student exposure and social experience so that they can come to grips effectively with important social and intellectual issues. In this chapter we are concerned with ways to assess the latter situations, and to plan for them. From a planning-assessment standpoint, these problems involve two stages of activity.

Identifying Gaps in Student Experience and Comprehension of Issues. One can use information gained through the organizational streaming analysis, just described, to identify expectable breaks in social experience that may need bridging if the intellectual concerns of a course are to come alive for students. Some students (usually the ones we like most to teach) are intellectually curious and enjoy thought worlds sufficiently to play with any ideas presented to them. Many other students, however, will take intellectual tasks seriously only after they have come to see the relevance of questions being asked to things they care about. For areas outside their previous experience, this often requires some kind of encounter that touches them at a level beyond intellectual description or analysis. What will do this varies enormously from student to student.

Some people become excited about analytic issues or about methodological questions when they sense the aesthetic issue of proof or rejection of claims that are involved. Many more, however, become absorbed in such issues only after they begin to care about knowing the answer to a question. For macro-level sociological questions, coming to care about the question often involves discovering that it has direct effect on one's own life, or on the lives of other people with whom one feels some identity--and that the relationships studied can work in more than one way.

There are many routes by which such awareness can be developed, of course, including well-designed reading assignments, films, simulations, field encounters, and data gathering that lets one see how social patterns affect other people's lives. Which particular route is used probably is less important than to plan some such opportunity as part of the course. We cannot assume that students will share our own sense of the importance of a question. They may have to discover its relevance, at a gut level, in order to care about the analysis presented in a course.

One can combine "streaming information" (as described earlier) with observation of past classroom experience as various topics are raised. Together, these help predict which topics of a course will be immediately salient to students and which may need special introduction. The salience of various macro-analytic questions tends to vary from student cohort to student cohort. It can be helpful, when planning, to remember how events and issues that occurred before one was born, or before reaching adolescence, feel at an emotional level. One then can ask how old these students were when social events occurred that might make a course topic relevant. This approach can help the instructor become aware of gaps in common social experience--or shifts in the nature of public issues--that may affect the gut level comprehension of analytic questions addressed in a course. While this may not affect which topics are addressed, it helps one realize the kinds of encounters that may have to be generated if the bulk of students taking the course are likely to become engaged in the question.

Mobilizing Teaching Resources Available through Students. It now becomes possible to use more fine-grained analysis of course populations to assess the easiest ways to provide entrée to problems or experiences that are foreign to identifiable groups of students. Students not only represent categories of social experience (e.g., members of the upper middle class, those with previous training in particular cognitive styles, social cohorts); each student also represents a unique constellation of experience and point of access to others outside the classroom. The impact of a course can be far greater if those class members who have unique experience or intellectual understandings can become points of access to such questions for other students. For some courses it is possible to predict which special skills or backgrounds are likely to be found almost any time the course is taught and to plan ahead of time to make use of such opportunities. For others, it involves advance planning alternatives--routes to follow if certain kinds of students appear on the scene, other ways to proceed if no one emerges with the qualities one would hope for.

Assessment, early in the term, of experiences, knowledge, and skills available within the particular population of a classroom, allows one to plan so that teaching becomes a two-way street. A fairly easy way to do this involves alerting students during the first week of class to the topics or areas of a course in which their contributions could be helpful. In large enrollment courses, for example, I have occasionally used a questionnaire during the first or second week to discover kinds of previous knowledge available within the student population, persons who have access to individuals or social settings that would be relevant to later topics addressed in the class, and those class members who would be willing to make their particular resources available to other members of the class.

III. Resources Students May Have for Assisting with the Teaching Process

Previous Social Experience. Large courses, quite difficult learning environments in many respects, provide one special resource: they are likely to include some individuals with unusual access to particular topics being studied, and unique vantage points for studying them. (For example, in my introductory sociology course this year--with eleven hundred students over two terms--we had the sons and daughters of union leaders, neighbors of major corporation executives, people who had worked on assembly lines in the auto industry, acquaintances of political machine personnel and city council members in the city government of a nearby metropolitan center, a woman with direct access to information about the inner workings of one of the major popular culture organizations, and the like. Topics that otherwise were foreign and artificial began to come alive as members of the class introduced us to the real personnel involved in some of the issues we were studying.) Normally we do not know who's out there in a large course. It is a fairly simple matter to survey the social networks and routes to which members of the course have access. (It is more complex, however, to devise a simple procedure for plugging these in with sufficient lead time to take advantage of them.)

Different Kinds of Skills. In addition to variations in experience and information, students vary in the skill levels they bring to a course. These include both intellectual and social skills that could be useful for helping other students learn. An early assessment of skill levels can lead to assignments geared not only to the abilities of different sets of students (so that all move forward during the term) but also to possibilities for letting students teach one another. The same problem or topic can be pursued with differing degrees of finesse, sophistication, or depth, if assignments are adjusted to the initial skill levels of students. Assignments that approach the same material at different levels can be useful teaching devices for the less advanced students, if feedback of student work is built into the course plan: they can see what *else* can be done with a question, once one has mastered the initial skills needed to approach it. If assignments also are designed to tap the creativity of the students, the work of beginners can be interesting and instructive to the more advanced students as well.[2]

Assignments in which more knowledgeable students cement their skills by helping less knowledgeable students acquire these tools can be useful to both groups and provide one-to-one teaching opportunity. The learning environment becomes most comfortable for all concerned, however, if you can generate assignments that let such peer teaching become a two-way street, so that today's learners become tomorrow's teachers, and vice versa. Such a learning exchange might trade information for skill, when different levels of achievement are involved, or might involve *social* skills rather than primarily verbal or mathematical intellectual skills; but it is a rare class that does not include people who have unique things that will be valuable for others to learn. The teaching task involves anticipating the kinds of differential abilities and knowledge likely to be present in a class and then designing ways to let these emerge, through assignments or through class interaction.

In short, approaching a student population as a cluster of resources that can be mobilized and shared, rather than as a group to be taught at some middle level at all times, often results in much more vital learning experiences for both teachers and students. An assessment procedure that anticipates the kinds of talents that would be useful for a particular course and systematically identifies these within the currently enrolled class members can allow this kind of teaching to emerge without a sense of hectic, daily planning: if the broad outlines have been anticipated in advance, the instructor can build segments of the course into the syllabus that are aimed at taking advantage of such differentiation among the students.

Creating a More Desirable Mix of Teaching Resources Among Students. Sometimes one may want to reschedule the time at which a class meets, in order to draw in a population group that would assist with the goals just described. If some courses, for example, were taught in the evening, they would draw a mix of adults and current students. Where a blend of ages or backgrounds would change the nature of classroom interaction, such rescheduling might be a simple course innovation. In addition to the blend of students likely to be enrolled in a course, teaching resources from outside may be differentially available depending on the time at which a course meets. (Such outside resources include visiting speakers, adjunct lecturers or co-teachers for a portion of the course, or class experiential assignments.)

IV. Planning to Use Different Kinds of Diversity Within the Student Population to Maximum Advantage for Teaching

Engaged learning becomes a different kind of challenge, depending on the type and degree of heterogeneity found among students in a particular class. Students vary in the social backgrounds and previous experience they bring to a class; they also vary in skills and knowledge, styles of intellectual work, and the kinds of interests they bring to a course. Previously I alluded to such differences and suggested that such diversity can be a valuable teaching resource. It also can create serious teaching problems, as many readers will know only too vividly from their own experience. It is time to face this issue head-on, suggesting various strategies for planning that take potential diversity into account. One can attempt either to minimize its impact or to maximize it, depending on the specific goals of the course and the instructor.

Diversity in Social Backgrounds. Heterogeneity of social background and interests can be approached in various ways. The large enrollment class, for example, can advertise a few special sections for persons with special interests, and can direct field experiences, research assignments, and supplementary readings to that special interest. (For example, in an introductory course that includes many nursing students, there could be sections dealing with health issues--or perhaps with nursing as a special role within the health care field; or an introductory course might include sections comparing how different societies deal with the practices studied in the course. There could be a sub-area focus, on Africa, for example, or on the Latino experience in the U.S.) Smaller enrollment courses can form work groups that meet together occasionally to pursue such topics or to prepare a report to share with the rest of the class. This approach seems to work best when themes that subgroups pursue have been identified initially and addressed within the class as a whole; it can become a rich experience for the entire group if the subgroup has some opportunity to feed back into the rest of the class what they have learned and how it affects their sense of the larger issues and arguments addressed in the course.

Widely ranging social backgrounds within a classroom can be a distinct teaching asset, of course, and many teachers prefer to keep the mix as varied as possible. Indeed, one of the few values that the large enrollment introductory sociology course offers pedagogically is the variety of previous experience and the range of potential access routes to various social settings that regularly emerge among the students enrolled in such a course.

To tap such heterogeneity in a way that illuminates rather than stereotypes various categories in the class requires special planning. Ideally, students from heterogenous social backgrounds should encounter one another interactively rather than simply through verbal exchanges in response to something the teacher said, so that the sharing becomes spontaneous and real for all parties.

For example, a jointly undertaken and discussed field observation project can let each person's knowledge and perceptions enrich what others understand about a setting. Students can learn to see different things, from what other students notice; they can begin to sense the context in which behaviors that seem strange to them make sense to the participants, if someone who has been there is helping interpret what is going on.

Similarly, data gathering and interpretation assignments which mix students of different backgrounds as team members can provide a natural means for students to bring their previous experience to bear on the puzzle of understanding what is going on in a particular social setting. If one wants to accomplish this, clearly not just any field project or data gathering assignment will work. Experiences need to be chosen so as that at least one student's background will provide relevant information for helping others understand a situation--what it feels like to be in that situation, what kinds of constraints lie behind the behavior choices being observed, how certain situations are managed by the participants, or something about the larger-scale dynamics creating this situation. Wiseman and Aron (1970) have developed a manual of field projects for sociology students that can help spark ideas adaptable to a variety of courses.

In addition to field observation and data gathering or interpretation projects, game simulations can accomplish something of the same thing, especially if they allow students to reverse structural situations temporarily to discover underlying emotional responses or motivations that arise in themselves when they have to deal with a different set of situational variables. Greenblat and Duke (1975) assess pedagogical issues involved in game simulation. Gillespie (1977) provides a helpful catalog of games useful for sociology classes.

Once differences in students' backgrounds have emerged naturally, it often becomes possible for students to share their situations unselfconsciously with one another. This is most likely to happen, however, if initial bases for trust or at least for interest in each other have had a chance to emerge. Naturally occurring interactions, if planned into a course, can let self- or situational-revelation occur without discomfort. (I cherish the response a black woman student once made to a request from other students to "tell us about the black experience." "I can't," she said, "its the only one I've ever known.")[3] A learning situation that lets students contrast what they see or experience with what they have known in the past is more likely to produce insightful, as contrasted with stereotypical, contributions from informants.

If the instructor builds time into a course for inputs from community resource people, he or she can use the social networks likely to be available within a large-enrollment class. Students can help to locate people or information accessible to some members of the class but not to the instructor personally. A simple questionnaire, indicating areas of experience or expertise that would be useful for this course and asking people to indicate persons they know or routes they could use to reach such people or settings, can alert the instructor to rich potential resources--provided time has been built into the course plan for following through on some of this material. How much more useful to use students' first-hand sense of the situation than to discuss questions abstractly without reference to real people.

In addition to alerting one to potential classroom resources, such network assessment in a large course can let one individualize assignments. Some students, for example, can check assertions being made by the instructor or by statistical data against impressions and information provided by people in such settings. Brought back to the class this could provide a useful confirmation, challenge, or interpretation of points made more abstractly.

Diversity of Skills. Skill level differences, in such things as previous background in statistics, methodological sophistication, or awareness of theories and concepts used in sociology can be identified early in the term. Students can be given a variety of self-pacing options for covering part of the course material. Several articles describe variations on this approach used in a variety of settings for teaching undergraduate sociology. These include TIPS (Bassis and Allen, 1977), PSI (personalized system of instruction) (Clark, 1974; Ludwig, 1975), SSI (self-scheduled introductory sociology) (Rogers, Sartariano and Rogers, 1977), and computer-based instructional programs (Conklin, 1976; Best, 1977).

Alternatively, students with advanced skills can be recruited as tutors to help those who are encountering such material for the first time. It frequently is a valuable experience for the tutors (who solidify their grasp on the material in question or come to understand it at a deeper level in the process of teaching it) as well as for the tutored. It also can help with the pacing of a class: if the less-advanced students have special attention given them by those already in command of some of the material, the whole class can move forward together. (Such tutoring arrangements require consultation and supervision for the tutoring to work well. And they seem to work best if the tutored can later help the tutors learn something new, so that status differentiation within the class is not divisive.)

Diverse Intellectual Styles. Heterogeneity of intellectual styles can be dealt with quite consciously when planning a course. One can include descriptive, analytic, and experiential material within the same class hour, for example, so that students who learn in different modes can pick up those aspects of a topic for which they are ready, and can have available differing levels or modes of analysis to use as they become personally accessible for other ways of learning. One can include items within a class session that appeal to the playful and the judgmental sides of a person as well as to the intellectually analytical elements of their mind. In short, one can plan class session ingredients so that a variety of learning styles have a chance to become engaged.

Diversity of Interests. Heterogeneity of interests is more difficult to deal with creatively, in the planning of a course. One can of course select a series of topics for attention that are likely to interest different streams of the expected student population. There is no guarantee, however, that the other students in the course will be interested in those topics, unless one plans the presentation in a way that lets the specially-interested students share their sense of excitement about that area. The teaching task involves finding a point of connection that will let a student discover an interest that has not yet emerged.

The planning task is similar to that encountered when planning for a homogenous student population that lacks certain kinds of experience or previous knowledge that can make an intellectual question real. Once again, tried and true formulas for many teachers involve generating exposures that will give these students a sense of the real situations and people who are affected by the issues being discussed. Films (if emotionally effective), field trips, sensitively written participant-observation studies; personal documents or even novels, can let students encounter concepts, arguments, contradictory data sources, or the like, at more than a verbal level, so that they begin to care about solving the intellectual puzzle being addressed.

Strategies available for course planning become much more focused once one has analyzed the population streams regularly coming into a particular course. The degree of homogeneity or heterogeneity of social experience, intellectual style, preparation, and previous exposure to sociological modes of analysis, can alert one to resources likely to be available among the students in a particular class. They also make one aware of the special needs likely to be encountered because of the population stream from which this course or university recruits and because of the organizational rules within the school that direct particular groups of students your way.

There is no automatic, right mix of backgrounds for students taking a course. Sub-grouping homogeneously allows one to plan parts of the course so that specialized interests can be addressed without boring the rest of the student population. Keeping heterogeneous grouping of social backgrounds allows one to use natural dynamics occurring within the student population. These can make real for students, at an experiential level, some of the macro-sociological forces that shape their identities and experience.

Ideally, one might want to take advantage of both options, by mixing the grouping of students who encounter one another over the course of a term. This probably works best in middle-sized classes (neither enormous, so that one groups students into sections that meet regularly together, nor too small to produce a predictable range of desired backgrounds available to be grouped and regrouped). Class size, of course, can be manipulated by setting enrollment ceilings and by timing how often a course is offered, so that a backlog of demand for it develops naturally. (Each class size has its own natural dynamics, however, and different teachers work comfortably with different sized groups.) Thus there is no correct solution to the size of heterogeneity/homogeneity of backgrounds in question; rather there is a series of inter-related problems and possibilities inherent in each mix. These should be taken into account when planning a course.

If class size and degree of heterogeneity is a given, one has a real, though limited, range of options for helping this subject matter come alive for students. Where class size and degree of heterogeneity of background can be verified through timing and recruiting practices, a wider range of planning options emerges.

Assessing the Organization of Learning Opportunities

A final area of assessment for planning does not fall easily within these categories, which have been aimed toward use by an individual instructor. Students often benefit from a mix of learning formats--some individualized instruction, some pre-programmed work-at-your-own-pace instruction, some genuinely independent study, some conventional lecturing, seminars, field experiences, and genuine in-depth research opportunity. Individual instructors rarely have a chance to think beyond some desirable mix in their own courses. Departments, however, can jointly assess the range of learning formats available, and faculty interest in experimenting with approaches other than those most in use.

If one thinks about credit-hour generation rather than about typical class size, it sometimes becomes possible to try quite varied formats for learning. In many departments teaching load is calculated by multiplying the number of credit hours per student times the number of students taking a course. Departments regularly pay for smaller enrollment courses by balancing them out with classes that have large enrollments. An additional strategy is available, however, a course that offers more credit hours per student, so that the faculty member generates the required number of credit hours while teaching fewer students. They in turn, put in proportionately more work for this particular course. (If bureaucratic red-tape makes increased credits for a course impractical, it sometimes is possible to have a limited number of students sign up for two course numbers taught by the same instructor and organized around a single intellectual focus for the two courses.) If the credit-hour-per-student-instructed increases for an individual course, the number of students that must be taught by an individual instructor decreases proportionately. Thus varying the number of credits available per course can be a simple way to provide variety in size of class, while still generating the required number of credit hours per instructor. (In the University of Michigan's Residential College, for example, we are experimenting with a team-taught 12-credit course involving two to three small research teams who do serious joint social research for a term--combining faculty teaching and research time, giving students in-depth exposure to research problems and methods of the social sciences. Joint publications are generated from these projects.)

Examples could multiply, one hopes, endlessly. The point, however, should by now be clear: assessment of the people we teach should extend to the kinds of formats needed for optimum learning by them, with assessment extended beyond the individual course to the larger curriculum and the options available within it.

Postscript. These comments were sketched primarily from the vantage-point of undergraduate teaching. From my observation post as one who teaches at all levels--from introductory courses through undergraduate and graduate courses, to Ph.D. dissertations--I am convinced that equivalent planning issues apply at all levels of teaching, including graduate instruction. In graduate courses the range of student abilities and interests usually covers a more narrow scope so that issues and strategies become both more specific and less amenable to major tampering with instructional modes. Eventually we should address these questions to graduate as well as undergraduate instruction.

Notes

1. This assessment of knowledge available, knowledge needed, regarding the implications of differeing student mixes for course planning, relies on three sources for inputs: discussions in the literature, observations of experimental approaches tried by colleagues or by myself over a twelve-year period, and reports, many of them not in written form, from teachers around the country. To limit this assessment to a literature review seemed unduly restrictive, given the range of projects accessible for observation and the relative paucity of written material in the area.

I would like to single out several colleagues at the University of Michigan, who have experimented with many of the procedures discussed in this report that lack literature citations. While I have learned much from many colleagues, I am particularly indebted to the following people for a number of "field trials" that have become consolidated and reworked for inclusion in this report: Mark Chesler, Jim Crowfoot, William Gamson, Zelda Gamson, Jane Heirich, R. E. Iwirtzt, Peter Langer, Anne Larimore, Steve Martin, Andrew Modigliani, Ruth Simmons, and Marilyn Young. I thank each of them for the variety of things I have learned by watching them at work.

2. Some of my colleagues do remarkable teaching, of the kind just described. Jane Heirich's classes first made me aware that it is possible to mix ability levels creatively, as described here.

3. My thanks to R. E. Iwirtzt for sharing this anecdote.

References

BASSIS, M. and J. ALLEN (1977) "Teaching introductory sociology with TIPS: an evaluation." Teaching Sociology 4: 141-154.

BEST, J. (1977) "The introductory sociology survey." Teaching Sociology 4: 271-276.

CLARK, S. (1974) "An innovation for introductory sociology." Teaching Sociology 1: 131-142.

CONKLIN, G. (1976) "Using the computer in introductory sociology." Teaching Sociology 4: 83-96.

GILLESPIE, P. (1977) Learning Through Simulation Games. New York: Paulist Press.

GREENBLAT, C. and R. DUKE (1975) Gaming-Simulation: Rationale, Design and Application. New York: John Wiley and Sons.

KARABEL, J. and A. ASTIN (1975) "Social class, academic ability, and college 'quality'." Social Forces 63: 361-398.

LUDWIG, M. (1976) "The PSI social science program at Cuyahoga community college." Community College Social Science Quarterly, V, 4 and VI. 1: 25-26, 29.

MEDSKER, L. and J. TRENT (1972) "Factors related to type of college attended," pp. 41-50 in K. Feldman (ed.) College and Student. New York: Pergamon Press, Inc.

ROGERS, S., SARTARIANO, W. and E. ROGERS (1977) "Self-scheduled introductory sociology: an evaluation." Teaching Sociology 4: 125-140.

WILSON, E. K. (1966) "The entering student: attributes and agents of change," pp. 71-106 in T. Newcomb and E. K. Wilson (eds.) College Peer Groups. Chicago: Aldine Publishing Company.

WISEMAN, J. and M. ARON (1970) Field Projects for Sociology Students. Cambridge, Mass.: Schenkman Publishing Company.

Finding Optimal Instructional Programs and Procedures for Individual Students

Daniel Solomon

U.S. Bureau of the Census

The problem to which this chapter is addressed is that of identifying those characteristics of students which help determine their benefits from different educational programs, approaches, styles, classroom conditions, or specific instructional procedures. The aim, to put it another way, is to find ways to identify optimal "matches" between students and instructional characteristics. Although a substantial amount of educational research has been devoted to this topic, none of it, apparently, has involved college sociology instruction. It should nevertheless be useful to examine this literature, as providing the best available clues to the types of relationships which may be expected when similar investigations are conducted in sociology instruction. Before discussing this educational research literature, however, I would like to place the issue within a more general framework by briefly considering similar issues which have been problems for social science in general.

The notion that different persons respond differently to different situations and settings seems nearly a truism. Yet such a notion has had a hard time finding its way theoretically and empirically in the social sciences. It has surfaced periodically, but has never been part of the main stream. It involves a concern with interactions, while the predominant tendency in social science seems to have been with main effects. The concentration has been on one or the other (persons or environments), or on the unilateral influence of one on the other, but generally not on interactive effects between the two.

Interactions between persons and environments formed a central element in Lewin's (1935) approach to the prediction of behavior, with subjective perceptions of the situations (e.g., "life space") as important mediating variables. A similar approach was taken by Murray (1938), whose theory posited a specific set of situational "presses" and a set of parallel personal "needs," predicting optimal effects in situations whose presses correspond to the needs of specific individuals. More recent approaches to personality theory (e.g., Rotter, 1954; Mischel, 1973) have also stressed interactions of persons with situations.

A relevant stream of research, beginning with the Hartshorne and May (1928) studies, has investigated the cross-situational generality of human behavior. In some of the more recent research in this area, Endler and Hunt (1969) demonstrated that neither persons nor situations considered separately accounted for as much variance as their various interactions. Other investigators have attempted to relate individual differences in cross-situational generality of behavior to perceptions of situations and, in particular, to perceptions of similarity between different situations (Bem and Allen, 1974; Solomon, 1977). The importance of subjective perceptions of situations has also been stressed by Coutu (1949) who spoke of subjective "meanings" and personal "selectors" as internal factors determining the individual's reaction to a situation. Kelly conceptualized "personal constructs" in a similar vein, while Allport (1937) related similarity of behavior across situations to subjectively defined "equivalence of stimuli." Mischel (1973) has suggested several cognitive social learning variables which may mediate cross-situational behavioral consistency.

Although these interactionist tendencies have existed for some time, and continue to inform theoretical and empirical work, the search for main effects and attempts to maintain intra-disciplinary integrity still reign supreme. Ekehammar (1974), in a recent review, categorized theories of human behavior into three types--"personological" (giving primacy to traits and similar factors). "situational" (giving primacy to situational factors), and "interactionist" (emphasizing person by situation interactions). His consideration of the research evidence led him to conclude that the interactionist position was the most clearly supported. Cronbach (1957) discussed the bifurcation of psychologists into "correlationists" and "experimentalists." The correlationists examine relationships between different individual characteristics, assume cross-situational generality of behavior, and consider variation between situations and/or interaction with situational characteristics "noise" to be avoided or ignored. The experimentalists manipulate situational parameters and aim at finding effects which hold across individuals; factors which are influenced by individual differences are discounted as not contributing to the development of general laws. Cronbach called for an effort to integrate these two approaches. More recently, Cartwright (1978) has called for a new emphasis on a truly integrative social psychology which focuses on the interactions between persons and society.

Research in education has also tended to be non-interactionist. One stream of research has attempted to identify instructional programs and procedures which would prove effective for students of all types, and has tended to ignore individual differences in reactions to the programs. Another stream has concentrated on identifying individual aptitudes, motives, etc. which best predict educational outcomes, independent of variations in instructional treatments--the emphasis here has been on identifying students most likely to succeed, and has often been used for selection purposes.

Although some research exploring interactions between individuals and educational situations has been conducted at least since the thirties, it has become fairly extensive in the last decade or so. A comprehensive methodological and substantive review and critique of the entire area has recently been published by Cronbach and Snow (1977). Other reviews have been produced by Berliner and Cahen (1973), Bracht (1970), Hunt (1971, 1975), and Tobias (1976). The label most commonly applied to this area of investigation is "aptitude-treatment interaction," although the phrases "trait-treatment interaction" and "attribute-treatment interaction" have also been used. All of these terms seem too limiting, although Cronbach and Snow (1977), who use the former, consider "aptitude" to be a general term referring to any individual characteristic which may help predict an outcome, and "treatment" to be any aspect of an educational program provided to students. Hunt (1975) prefers the term "person-environment interaction"; it is more inclusive, and allows for the consideration of interactive effects for total persons (rather than, or in addition to, abstracted attributes of persons), and of total environments or situations (thus, an investigation could include the school atmosphere, in addition to the classroom, as part of the impinging "environment"). Although the phrase, "aptitude-treatment interaction" (ATI) will be used in this chapter, it should be understood in a general and inclusive sense.

Most of the research which has been reported has involved short-term experiments, usually manipulating one or two treatment characteristics and examining interactions with one or two person attributes (e.g., anxiety, intelligence). A smaller body of research has involved field experiments, in which experimental treatments are distributed among two or more teachers or classrooms, students (in some cases) are assigned randomly to the different treatments, and interactions between the treatments and measured student characteristics are examined for their effects on measured outcomes. A large investigation by Goldberg (1969) is a good example of this approach. Another research approach which has been applied to this problem is a completely natural-state approach. In this approach, the least frequently used, characteristics of existing educational settings are determined (through, for example, observations, and student questionnaires), measures of student characteristics are obtained, and the effects of interactions (and main effects) on educational outcomes are determined. This approach has been used by Solomon and Kendall (1976). There are advantages and disadvantages to each approach. The experimental approach affords greater control over the investigated variables, but tends to use fairly artificial situations, investigates short-term effects which may have little relationship to actual educational experiences, and tends to treat variables in artificial isolation rather than the multivariate complexity of the natural educational setting. Natural state studies provide no experimental controls, have difficulty in establishing causal linkages, and are beset by confounded variables; yet they do approach the educational situation as it is experienced and acted upon by teachers and students. It seems clear that there is a need for each of these approaches; the deficiencies of each can be weighed and accounted for by comparing results across them. Findings which recur with different approaches are the ones which should inspire the greatest confidence.

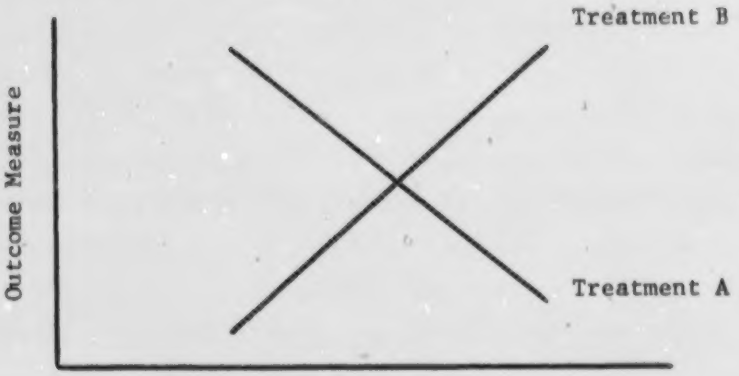
Various statistical procedures have also been applied to ATI investigations. The most common have been analysis of variance and regression analysis. Typically, when analysis of variance is used, the person variables are blocked, or upper and lower extreme groups are used, with other subjects eliminated from the analyses. Cronbach and Snow (1977) criticize the blocking procedure because of the amount of information lost. They also feel that the extreme-groups procedure is to be recommended over that of blocking the entire distribution, because it has greater power. They generally recommend that an N of 100 per treatment cell is required to provide adequate power to an ATI experiment.

Cronbach and Snow suggest that regression analysis is usually the preferred method of analysis. It makes use of the maximum information contained in the individual attribute measures. The typical procedure would be to compare the regression of outcome measure on attributes in each of the experimental treatments. If the regression lines differ significantly between treatments, there is evidence of an attribute-treatment interaction. There has been some discussion about the distinction between, and relative value of "ordinal" and "disordinal" interactions (see Figure 1). A "disordinal" interaction is one in which the regression lines for two treatments cross--this would indicate, for example, that students scoring low on attribute X perform better in treatment A than treatment B, while students scoring high on the same attribute perform better in treatment B than treatment A. An "ordinal" interaction, on the other hand, is one in which the regression lines are not parallel, but do not cross--thus, students scoring low on attribute X might score slightly higher in treatment A than treatment B, while students scoring high on the same attribute might score much higher in treatment A than in treatment B (of course, many other ordinal and disordinal interaction forms are possible). Bracht (1970) considered only disordinal interactions to be of theoretical and practical interest. Others (e.g., Cronbach and Snow, 1977; Berliner and Cahen, 1973) feel that both types of interaction are worthy of interest and attention.

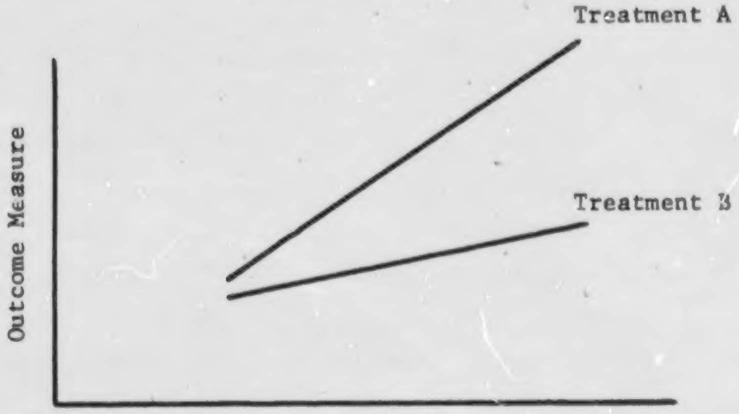
The above examples include a single attribute and a single treatment. Experiments have been conducted which include several variables of each type. This adds complexity but does not alter the basic logic of the approach. It is also possible to examine higher-order interactions (i.e., investigating the joint effects of two or more person variables and two or more treatment aspects). Thus, in a recent example, Peterson (1977) obtained a four-way interaction between student ability, student anxiety, structuredness of instruction, and amount of class participation.

An extension of the regression procedure is to incorporate interactions into the regression analyses as product terms (Kerlinger and Pedhazur, 1973). This is possible both with experiments (in which the treatments are represented as "dummy" variables), and with natural-state studies, involving measurement of large numbers of continuous variables representing aspects of persons and of situations. The latter procedure would also lend itself well to path analyses.

One of the most perplexing methodological problems in ATI research has been that of the appropriate unit of analysis. Most of the experiments which have been reported, until recently, used the individual student as the unit of analysis. This would be appropriate only if each subject received a "treatment" independently of other subjects. But when the students are organized into ongoing classrooms (or other settings), the treatment (or


Treatment B
Treatment A
Outcome Measure
Attribute X

a. Disordinal Interaction


Treatment A
Treatment B
Outcome Measure
Attribute X

b. Ordinal Interaction

Figure 1: HYPOTHETICAL ILLUSTRATION OF "ORDINAL" AND "DISORDINAL" INTERACTIONS.

educational experience) of one student cannot be considered to be independent of that of any of the others in the same classroom. Treating the student as the unit of analysis would clearly overestimate any classroom variable main effects, as well as any interactions involving classrooms. Yet, the investigation of student by classroom interactions would seem to be precluded by using the classroom as the unit of analysis. Two approaches to this problem have been suggested. Cronbach and Snow (1977) advocate conducting regression analyses at several levels (e.g., within classroom, within treatment) and comparing the results obtained across these levels. Page (1975) has recommended a repeated-measures analysis of variance approach, with the classroom treated as the unit of analysis, and means for different subgroups within the class (e.g., students low on anxiety, students high on anxiety) as "repeated measures" of the classroom. The Page approach was applied in a recent study by Solomon and Kendall (1976).

In spite of the fact that many ATI studies have been done, few solidly established and validated generalizations can as yet be made. Cronbach and Snow (1977) have conducted the most extensive survey of this literature, including careful examination of the methods and data-analytic procedures, and, in some cases, re-analysis of the data. They have summarized the trends which they feel are well-established in this body of research (as of about 1975). The following paragraphs are largely based on the Cronbach and Snow summary.

Among the individual characteristics included in ATI research have been general abilities, specialized abilities, and personality traits or styles. Some of the clearest findings have occurred with general abilities. Generally, high ability students appear to benefit from the opportunity to organize and process information themselves, while low ability students do better if much of the organizing is presented to them. Among other studies, strong support for this generalization has been provided by a program of research, that students with a high "conceptual level" (referring both to cognitive complexity and interpersonal maturity) perform best with low structure, or are unaffected by variations in structure, while those with low "conceptual level" are benefited by greater degrees of structure (e.g., operationalized, in different studies, as including lecturing, presenting rules before examples). Apparently, high ability students benefit from settings which require their own intellectual performance, and low ability students from those in which there are no such requirements, or the intellectual content is provided to the students directly.

Among the specialized abilities which have been investigated in ATI research are spatial and mathematical abilities, memory, and prior learning experience. Of these, prior learning experience has shown the clearest interactions with instructional variables. Students who have had prior experience with a particular treatment (usually a specialized procedure of some sort) are most likely to perform well under that treatment.

A number of studies have looked at interactions with various personality characteristics. The general variable which has shown the greatest evidence of interactive effects is called "constructive motivation" by Cronbach and Snow. These interactions have indicated, in general, that constructively motivated students perform best in classroom situations which are less directive, more spontaneous, and require students to take more active roles. Student anxiety has also been included in a number of studies. Some studies have indicated that anxious students do better when the treatment makes fewer demands for self-direction or independent thinking. There is also some evidence that anxiety interacts with stressfulness of treatments.

Some studies have indicated that student need for affiliation interacts with supportiveness of the atmosphere and with the opportunity for interaction, to influence student satisfaction. Studies by McKeachie (1961) and by French (1958) show, respectively, that the type of "cues" provided to the subject about the nature of the situation (achievement-relevant, affiliation-relevant, power-relevant), and the type of "feedback" given to subjects about performance ("feeling feedback" vs. "task feedback") determined the degree to which individuals' personality needs (for achievement, affiliation, power) were related to their performance. This research focused on those situational factors or cues which were expected to activate or capitalize on specific relevant student needs--thus it was an inherently interactive approach, akin to the "need-press" pairings of Murray, the Hunt approach, and others.

The above summary of research findings includes studies done at all academic levels. Some of the results may generalize across levels; others may not. In particular, the degree to which these findings apply to college sociology instruction is unknown, and unknowable without a substantial new research effort. However, most of the above generalizations are based on groups of studies, at least some of which involved college students. (More so for the studies involving personality than those involving measures of ability.) Therefore, unless sociology students are radically different from other students, it seems likely that the results will prove applicable to them as well. Furthermore, given the present state of knowledge, it is probably more reasonable to base practical decisions and practices on whatever empirical evidence is available, shaky though it may be, than on none at all.

Still, it would be very difficult, and probably unwise, to generate specific prescriptions for approaches to college sociology instruction from the above set of generalizations. It would probably be more appropriate to consider these to be suggestive findings which instructors should be aware of, so that they may conduct their own informal assessments of applicability to particular instructional settings, and make their own judgments about the validity and utility of these generalizations when actually applied in those settings. These findings may also suggest other consistent approaches, not specifically investigated in the prior research, which may have greater relevance to college sociology instructions. It is to be hoped that research applying some of these approaches to the investigation of college sociology instruction will also be undertaken, and that, if such research efforts are instituted, they will profit from, and improve upon, the examples found in prior research.

There is a great need for additional research in this area at all educational levels and with all types of educational content. There are shortcomings with much of the prior research which may account for the relative paucity of interactive results. Methodological problems and data-analytic problems have been extensively treated by Cronbach and Snow. An additional problem, noted by Goldberg (1969), has also yet to be remedied. He suggested that the failure of the personality measures which he used to more clearly interact with instructional treatment variables may have been due to the fact that such measures are deliberately constructed not to interact with situation. Personality measures derive, for the most part, from personality theories which assume cross-situational generality. The measures are therefore designed to be consistent across situations. The same is true of measures of achievement and aptitude, which are designed to rank people intellectually, independent of situation. Measures of person variables need to be designed specifically for the purpose of engaging in interactions with situation variables; the occurrence of such interactions could constitute one criterion for determining the validity of such measures. The measure of "conceptual level" (Hunt, 1971, 1975) was developed within a specifically interactive theory, and has provided one of the more consistent and convincing bodies of ATI results.

The ecological validity of much of the prior research is also open to serious question. Experiments which include only one or two person variables and one or two treatment variables may be well-designed to contribute to a small-scale theory, but they give little information about the operation of these variables in natural settings. At the least, they require replication within those settings. But when such replications are attempted, additional variables should also be investigated, since they exist in the settings and are undoubtedly also operating. Until one can treat a natural situation in something approaching its natural complexity, it may be impossible to make reasonable applications of research results. Some of the "natural state" studies are open to the same criticism. To the extent that such a study focuses only on a limited set of variables (ignoring all other possibly relevant ones), or includes many variables but examines their correlations or interaction effects one at a time or pair-by-pair, the true complexity of the situation is being ignored or misrepresented. Numerous multivariate statistical methods are available for handling this complexity and should be applied in such instances.

Further work toward specifying the types of functions different treatments can have with respect to different attributes also is needed. One useful framework for this has been provided by Salomon (1972). He describes three basic "heuristic models," called "remedial," "compensatory," and "preferential." The "remedial" model predicts optimal results when an educational program focuses on teaching an individual prerequisite skills in which he is deficient. The "compensatory" model focuses on treatments which bypass the student's deficiencies either by supplying external substitutes for them or by circumventing the need for them altogether through changing the situation so that the lacking skills, etc. will not be required. The "preferential" model attempts to "match" the student's skills and/or motives, to provide a setting which capitalizes on his strengths and inclinations. Results obtained by Solomon and Kendall (1976) suggest a fourth model or function, which might be called "nonpreferential" or "broadening."

There was evidence that some students benefited from being in situations directly opposed to their preferences. Children with expressed preferences for autonomy and self-expression performed best in controlled and orderly classrooms, while those whose stated preference was for more structure showed greater gains in classrooms which allowed more student autonomy and self-direction. The general explanation offered was that children at each extreme benefited from a setting which required them to experience a mode of activity which they would otherwise avoid, thus providing them with something which they lacked (greater discipline in their approach to tasks for the autonomous/expressive children; greater experience with choice, freedom, and self-direction for those more oriented toward external control and structure).

The preponderance of research on ATI has represented students by measures of ability, achievement or motivation, and "treatments" by measures or manipulations of instructional programs or teacher practices. Large areas of potential influence have been neglected. For example, there has been little examination of the degree to which the treatment is defined by the general institutional context and characteristics, by the classroom social structure, by contextual classroom variables (such as class size, sex ratio, ability distribution, and the like). Some of these have been looked at in partial ways by some investigators (for example, Stern, 1969, has devoted much attention to the identification of dimensions of institutional climates), but they have not as yet been incorporated into the general implicit ATI paradigm. Similarly, there has been relatively little investigation of the possible role of student perceptions of treatment characteristics and student classroom behaviors as factors which may be essential for mediating the joint effects of treatments and students. (One interesting example of such a function has been provided in research by Dowaliby and Schumer, 1973. They conducted two studies; in the first, "dominant" and "participative" teaching styles were manipulated and found to interact with student anxiety--anxious students doing better in the teacher-dominated class; in the second, the two styles were blended into a single treatment and students were asked for their perceptions of the teaching style. These perceptions interacted with student anxiety in the same way as had the manipulated styles in the first study.)

Further progress in this area would be aided by the development of a general framework laying out the levels and types of variables likely to be important in producing student-environment interactions. Following is a step toward the development of such a framework.

It seems reasonable to suggest four general categories of variables: environmental (including institutional characteristics, classroom characteristics, teaching characteristics and styles), student status (including cognitive characteristics, motivational characteristics), student responses (including student perceptions of environments and treatments, and student classroom behaviors and other course-related behaviors), and outcomes (including acquisition of factual content, general comprehension, development of interest).

Various types of interactive and direct relationships are possible and plausible within this general group of categories. These should be explored. For example, there may be interactions between institutional characteristics and classroom atmosphere characteristics. A given classroom style (e.g., rigid and controlled) may have a different effect on students in the context of different institutions (e.g., a similarly rigid and controlled institution vs. a more open and flexible one). Student cognitive and motivational factors may also interact (some of the prior research has demonstrated interactive effects of student anxiety and ability level, for example). And the student status categories, separately or in combination, may have interactive effects with each or all of the environmental categories. In addition, each of these environmental and student status categories may produce main effects on student responses and outcomes. These general relationships between categories of variables are shown in Figure 2. (Possible interactions between subcategories within the general categories are not shown, to minimize the complexity).

Generally, the student response categories are seen as playing a mediating role. It is expected that the interactions and main effects of the various environmental and student status categories will produce their immediate effects on student perceptions and behaviors, and that these will then directly influence the various outcomes. These response variables may also exert a reciprocal influence on the classroom atmosphere and teacher behavior aspects of the environmental category. A classroom is interactive in a behavioral as well as a statistical sense; yet research has usually treated influence as unidirectional--e.g., from teacher to student.

Each of the suggested general categories can encompass large numbers of component variables. Following is a partial list of those which seem

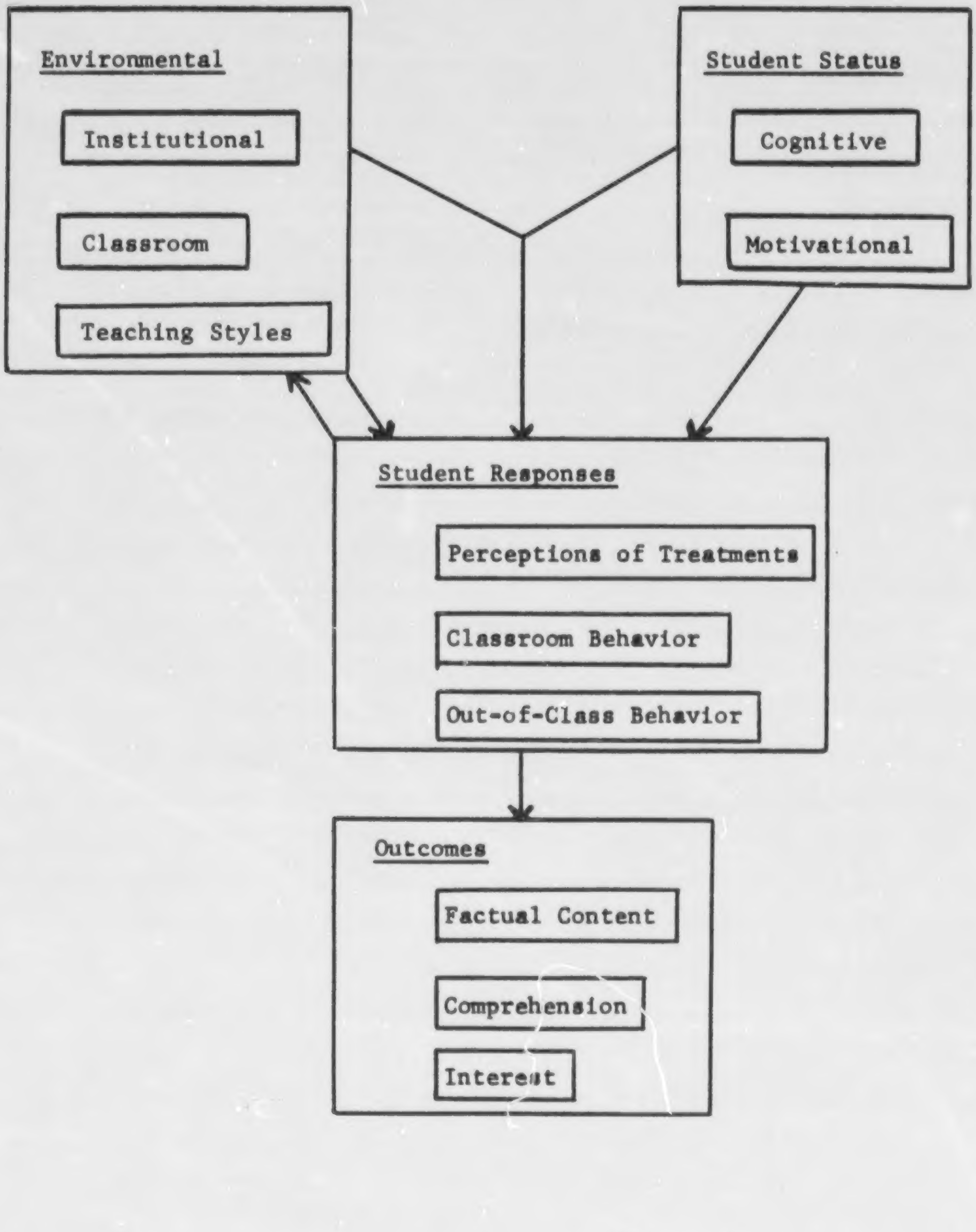

Environmental
Institutional
Classroom
Teaching Styles
Student Status
Cognitive
Motivational
Student Responses
Perceptions of Treatments
Classroom Behavior
Out-of-Class Behavior
Outcomes
Factual Content
Comprehension
Interest

Figure 2: SCHEME SHOWING INFLUENCES ON EDUCATIONAL OUTCOMES.

likely to be important within a general approach to research on aptitude-treatment interactions.

- A. Environmental
  - 1. Institutional
    - a. atmosphere
    - b. norms
    - c. population size
    - d. aggregate population characteristics (fields, affluence, goals, etc.)
  - 2. Classroom Characteristics
    - a. instructional method/teaching style
      - 1. lecture vs. discussion
      - 2. emphasis on transmission of facts vs. general principles
      - 3. inductive vs. deductive approach
      - 4. requiring student activity vs. student passivity
      - 5. degree of organization provided for student
      - 6. permissiveness vs. control
      - 7. warmth vs. coldness
      - 8. dryness vs. flamboyance
      - 9. direct vs. indirect teaching
      - 10. clarity of communication
      - 11. teacher mastery of material
    - b. classroom atmospnere
      - 1. friendly vs. aloof
      - 2. cooperative vs. competitive
      - 3. calm vs. excited
      - 4. stable vs. flexible
      - 5. broad vs. narrow norms
      - 6. role performance patterns (students and instructor)
      - 7. degree of hierarchalization in classroom role structure
      - 8. contextual characteristics of classroom group (ability, achievement, sex, age, race, class size, etc.)
- B. Student Status
  - 1. Student cognitive factors
    - a. general ability (conceptual level)
    - b. specific abilities
    - c. prior knowledge of subject

2. Student motivation/personality
    a. long-range goals
    b. short-range goals
    c. academic motivation
    d. preference for challenging tasks
    e. avoidance of failure
    f. conformity, orientation toward authority
    g. anxiety
    h. preference for structure
    i. preference for autonomy, self-direction
    j. self-esteem
    k. sense of control over environment
    l. preference for maintaining control
    m. preference for passive-receptive vs. active-seeking role

C. Student Responses
    1. Student perceptions
        a. of course requirements
        b. of instructor's expectations and requirements
        c. of teaching style/role
        d. of classroom atmosphere
        e. of position in class social structure
        f. of role in class
        g. of class norms and institution norms
    2. Student behaviors
        a. out-of-class discussions
        b. out-of-class studying/thinking
        c. specific classroom activities:
            1. listening
            2. note-taking
            3. questioning
            4. commenting
            5. discussing
            6. arguing
            7. participating in subgroup meetings

D. Outcomes
   1. Short-range (subdivision of course--lecture, week, month, unit, etc.)
      a. acquisition of information
      b. acquisition of principles
      c. stimulation of interest/motivation
   2. Medium-range (end of course)
      a. acquisition of information
      b. acquisition of principles
      c. stimulation of interest/motivation (taking further courses, relevant reading, etc.)
   3. Long-range (beyond end of course)
      a. retention of information
      b. retention of principles
      c. continued interest in field

More complete development of a listing of the important variables and dimensions in this domain must await further empirical work. In some areas (teacher behavior, classroom atmosphere, student ability, student motivation) previous conceptual and empirical efforts have identified certain global dimensions which seem to recur. These are reflected in the above list. Factor-analytic studies of teacher behavior, for example, have frequently found factors which seem to represent teacher warmth, permissiveness vs. control, and encouragement of participation (Solomon and Kendall, 1976). In other areas, however (e.g., student perceptions), little work has been done to date.

This set of categories or variables is intended as a step toward a comprehensive framework which could be used to guide theoretical development and research activity within this general area. It is not expected that any single study would encompass all, or even many of these levels. The purpose of developing a framework such as this is to make it possible to place the various research efforts within a common, general context, and to help identify sets of potential relationships needing further systematic exploration.

When a set of variables within the categories listed above (or a parallel set) has been developed, several strategies for continuing research will be possible. Large-scale multivariate studies might be conducted which attempt to deal with a good many of the variables simultaneously.

These would probably be "natural state" studies and would require large numbers of classrooms. One feasible procedure would be to factor analyze the variables from several general categories, and then to enter factor scores from the resulting factors into a series of analyses (probably multiple regression and path analyses), comparing main effects and interactions within and between classrooms. Another approach would be to select a smaller number of theoretically interesting variables and to investigate their interactions more intensively, either through a smaller-scale and more qualitative observational study, or through a tightly-controlled experiment. These strategies are not mutually exclusive. In fact, as was suggested earlier, the optimal strategy would probably be a combination of these approaches, either simultaneously, or building on each others' results, sequentially.

Another approach to the investigation of interactions has been explored recently--the use of cluster analysis. Cluster analysis (Blashfield, 1976) is a procedure for grouping cases according to similarity of profiles, and thereby identifying "types" which are internally consistent, and distinct from one another. It is considered to be useful for the empirical determination of taxonomies. The attractiveness of cluster analysis for ATI research is that it allows one to retain the complexity of the natural situation. It provides a way for looking at the effects on various educational outcomes of entities representing students and classrooms (or environments) in their natural groupings and with much of their natural complexity retained. This could prove to be a useful supplement to (and perhaps eventually even a replacement for) the more typical approaches which may analyze one or two isolated or abstracted dimensions at a time, looking at their main effects and interactions, while ignoring or attempting to hold constant statistically the simultaneous effects of other significant dimensions. The approach based on natural groupings accepts the complex of dimensions represented in a group profile, and looks at its total effects compared with those of other groupings. While intuitively this approach seems more likely to accurately represent social (and educational) reality, which _is_ complex, involves multiple simultaneous influences from numerous sources, and (perhaps) actually does form limited numbers of constellations of attributes (i.e., natural groupings), whether it will actually provide for greater theoretical development and greater usefulness in attempts at practical applications has yet to be demonstrated. An approach using total profiles allows for relatively simple analyses (once the clusters are established) and allows inferences to be drawn to total types comprising specific combinations of attributes. It does assume, however, that the identified types are truly coherent and internally consistent; if they are not, the procedure probably should not be used.

Most of the social science applications of cluster analysis have been devoted to the development of psychiatric taxonomies (Overall and Klett, 1972). A few recent attempts to identify types of students have also been made. In a study by Solomon and Kendall (1976), three student types were identified--low prior achievers with low self-esteem; self-confident, motivated, high prior achievers; and children who value autonomy and self-direction and are moderate achievers. Cunningham (1975) found four child clusters; the first contained high achieving, competent and advantaged children; the second, moderately high achievers who were also extroverted and cooperative; the third, low-achieving, introverted children, and the fourth, low-achieving, alientated and disruptive children. Bennett (1976) clustered children in terms of their personality profiles and produced eight clusters, including "unmotivated introverts," "unmotivated extroverts," "stable extrovert motivated conformists," "motivated stable extroverts," and "contentious nonconformists." All of the above applications of cluster analysis involved primary grade students. Applications to secondary or post-secondary students have not as yet been made, to my knowledge.

Although each of the above studies also applied cluster analysis to the identification of classroom types, the method has, in general, been used less in analyses of situational or environmental types than of person types. Frederickson (1972) called for efforts toward the development of a taxonomy of situations, and recommended factor analytic studies to identify important situational dimensions, followed by cluster analyses to identify situational types characterized by their profiles on the various dimensions. This recommendation has not as yet been widely followed, but was incorporated, in varying degrees, in the three studies cited above. Thus Solomon and Kendall (1976) identified six classroom types ("permissive and uncontrolled, with much student autonomy," "hostile, arbitrary and regimented," "controlled, disciplined, academically oriented, and supportive," "warm, oriented toward student expressiveness and creativity," "academically oriented, individualized interaction," and "controlled, non-individualized, impersonal"). Cunningham identified four (based on teachers' beliefs about instructional strategy rather than objective observations)--one combined a strong belief in teacher control with high scores on "subject matter integration" and very low scores on "teacher empathy" and "student direction." A second cluster showed low scores on "subject-centeredness" and moderate positions on all other dimensions. A third combined "student-centeredness" and low "teacher control" and "subject integration" with high scores on "teacher empathy" and "student direction." The fourth cluster was high on "subject-centeredness" but low on "teacher control." Bennett (1976) also clustered classrooms on the basis of teacher questionnaire responses. Twelve clusters resulted, representing various combinations of teacher control and student autonomy. Moos (1978) found five clusters of high- and junior-high school classrooms "control oriented," "innovation oriented," "affiliation oriented," "task oriented," and "competition oriented."

It is not likely that a complete set of "universal" types of social environments or classroom environments will be found. It is more plausible to expect that there are certain fairly common central dimensions of classroom life (and, more generally, of human group life), and that there are probably certain recurring patterns of combinations of these dimensions. Any single study, involving a particular sample of environments, will probably identify some "types" which closely resemble those found with other samples, and some which will be more limited to that sample alone. Only after comparing the results of numerous studies using similar methodologies but varying samples of environments will it be possible to say with some certainty which are the general, recurring "types" and which the more sample-specific ones.

There are at least two ways in which cluster analysis could be applied to the general categories identified earlier (environment, student status, student responses, and outcomes). Separate cluster analyses could be conducted with each of the first three general categories, investigating their effects on the outcome category by means of analysis of variance or generalized regression analysis. Or, subcategories within these general categories could be separately cluster analyzed (e.g., institutions and classrooms), so that more complex patternings could be investigated. A cluster analysis representing the total "environment" dimension could include factors representing the institution, the classroom atmosphere, teacher behavior dimensions, and classroom contextual variables. Each cluster would constitute an average profile representing positions on dimensions from all these levels.

Most of the above comments assume that the "environment" or "treatment" and the students are fairly stable for the duration of a course. It would be very valuable to explore the effects of certain changes and reciprocal effects within this context. Investigations of the degree to which teaching style and classroom atmosphere change during a course, factors producing such changes, the influence of the course characteristics and activities on student "status" characteristics, and the influence of these changes on interactions with "treatments" all would be very valuable to conduct. It might be possible to conduct time-series analyses, and to analyze interaction effects at various points during a course, not just at the end. Such efforts would of course add greatly to the complexity of such research efforts, but would seem to have potential for approaching a more realistic understanding of classroom dynamics.

All of the above procedures can also be applied to the investigations of interactions between student characteristics and ways of presenting various short sequences within a course--a brief sequence devoted to simulations, for example. Outcomes could be examined immediately after the conclusion of the sequence, and at the end of the course (and later) as well.

The implications of ATI research for courses containing heterogeneous groupings of students vs. those courses (or sections) in which there is less inter-student variability also need to be addressed. While it may often be possible to assign students to (or allow them to select) distinctly different course styles to match their needs and preferences, there may also be instances or settings in which this is unfeasible. There may, furthermore, be a positive value in maintaining student diversity within courses, particularly in some content areas, or for some topics. Although little research has focused on this, it may be possible to design courses so that students break up into groups for instruction on certain topics (which is then provided in different styles according to the needs of the students in the different groups), later reconvening into the total class group for other activities. Or, instruction may be provided on a more "individualized" basis, where students work at distinct tasks with periodic individual consultations with the instructor. Models for each of these exist in elementary education, but neither has been explored in the context of ATI research (with the exception of programmed instruction, which can be seen as a form of individualized instruction, and engendered a large body of ATI research in the 1960s). Other approaches which would allow instructors to capitalize on the findings of ATI research without setting up distinct and homogeneous courses or sections are undoubtedly also possible.

Research on aptitude-treatment interactions is still in an undeveloped state. It is difficult to do, has great methodological and statistical problems, and, as yet, has not produced a body of confirmed findings commensurate with the amount of effort which has been devoted to work in the area. Enough findings have accumulated to at least confirm the feeling that such interactions are real, however, and warrant the search.

References

ALLPORT, G. (1937) Personality: A Psychological Interpretation. New York: Holt.

BEM, D. and A. ALLEN (1974) "On predicting some of the people some of the time: the search for cross-situational consistencies in behavior." Psychological Review 81: 506-520.

BENNETT, N. (1976) Teaching Styles and Pupil Progress. Cambridge: Harvard University Press.

BERLINER, D. and L. CAHEN (1973) "Trait-treatment interaction and learning," pp. 58-94 in F. Kerlinger (ed.) Review of Research in Education, 1. Itsaca, Illinois: Peacock.

BLASHFIELD, R. (1976) "Mixture model tests of cluster analysis: accuracy of four hierarchical agglomerative methods. Psychological Bulletin 83: 377-388.

BRACHT, G. (1970) "Experimental factors related to aptitude-treatment interactions. Review of Educational Research 40: 627-645.

CARTWRIGHT, D. (1978) "Contemporary social psychology seen from a historical perspective." University of Michigan: Katz-Newcomb Lecture.

COUTU, W. (1949) Emergent Human Nature. New York: Knopf.

CRONBACH, L. (1957) "The two disciplines of scientific psychology." American Psychologist 12: 671-684.

CRONBACH, L. and R. SNOW (1977) Aptitudes and Instructional Methods. New York: Irvington.

CUNNINGHAM, W. (1975) "The impact of student-teacher pairings on teacher effectiveness." American Educational Research Journal 12: 169-189.

DOWALIBY, F. and H. SCHUMER (1973) "Teacher-centered vs. student-centered mode of college classroom instruction as related to manifest anxiety." Journal of Educational Psychology 64: 125-132.

EKEHAMMAR, B. (1974) "Interactionism in personality from a historical perspective." Psychological Bulletin 81: 1026-1048.

ENDLER, N. and J. HUNT (1969) "Generalizability of contributions from sources of variance in the S-R inventories of anxiousness." Journal of Personality 37: 1-24.

FREDERICKSEN, N. (1972) "Toward a taxonomy of situations." American Psychologist 27: 114-123.

FRENCH, E. (1958) "Effect of the interaction of motivation and feedback in task performance," in J. Atkinson (ed.) Motives in Fantasy, Action and Society. New York: Van Nostrand.

GOLDBERG, L. (1969) Student Personality Characteristics and Optimal College Learning Conditions. ORI Research Monograph, Volume 9, No. 1. Oregon Research Institute.

HARTSHORNE, H. and M. MAY (1928) Studies in the Nature of Character. Volume 1: Studies in Deceit. New York: Macmillan.

HUNT, D. (1975) "Person-environment interaction: a challenge found wanting before it was tried." Review of Educational Research 45: 209-230.

KELLY, G. (1955) The Psychology of Personal Constructs. New York: Norton.

KERLINGER, F. and E. PEDHAZUR (1973) Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston.

LEWIN, K. (1935) Dynamic Theory of Personality. New York: McGraw-Hill.

MCKEACHIE, W. (1961) "Motivation, teaching methods, and college learning," in M. Jones (ed.) Current Theory and Research in Motivation. Lincoln: University of Nebraska Press.

MISCHEL, W. (1973) "Toward a cognitive social learning reconceptualization of personality." Psychological Review 80: 252-283.

MOOS, R. (1978) "A typology of junior high and high school classrooms." American Educational Research Journal 15: 53-66.

OVERALL, J. and C. KLETT (1972) Applied Multivariate Analysis. New York: McGraw-Hill.

PAGE, E. (1975) "Statistically recapturing the richness within the classroom." Psychology in the Schools 12: 339-344.

PETERSON, P. (1977) "Interactive effects of student anxiety, achievement orientation, and teacher behavior on student achievement and attitude." Journal of Educational Psychology 69: 779-792.

ROTTER, J. (1954) Social Learning and Clinical Psychology. Englewood Cliffs: Prentice-Hall.

SALOMON, G. (1972) "Heuristic models for the generation of aptitude-treatment interaction hypotheses." Review of Educational Research 42: 327-343.

SOLOMON, D. (1977) "Perceptions of similarity between striving tasks and the generality of task preferences." Motivation and Emotion 1: 181-192.

SOLOMON, D. and A. KENDALL (1976) "Individual characteristics and children's performance in varied educational settings." Final Report, Spencer Foundation Project, Montgomery County Public Schools, Rockville, Maryland.

STERN, G. (1969) People in contexts. New York: Wiley.

TOBIAS, S. (1976) "Achievement treatment interactions." Review of Educational Research 46: 61-74.

Learning the Sociology of the Classroom

Thomas J. Rice
Denison University
Ruth Harriet Jacobs
Boston University
David A. Karp
Boston College

Sociologists have paid more attention to teaching in recent years, especially since the Project on Teaching Undergraduate Sociology began in 1974. The bibliography published by the Project (Davis, Goldsmid, and Perry, 1977) has revealed useful resources for many aspects of the teaching process, ranging from techniques for the introductory course to the selection of general textbooks. Yet we remain in the dark on the significance of classroom interaction and the host of variables which make for a particular level of performance (learning outcomes) under specified conditions. This brings us to our question: Why is it that some classroom contexts produce high levels of learning while others do not?

We currently know in a general way that some professors are more effective than others and that some specialized techniques (e.g., Simulations, PSI, T.A.s) seem to create some interest and even enthusiasm. However, we do not know the social processes by which these successes come about and how they may be replicated by those who have not yet come up with the magical mix of ingredients. Even introductory students can tell us that the classroom is a process involving complex interaction. But what is needed is a step beyond this elementary wisdom to transform it into concrete diagnoses and teaching operations. But before pursuing this task, let us elaborate on some of the more subtle problems of the classroom as social space.

The Classroom: Subsurface Web of Thought and Feeling

What is insufficiently realized in the analysis of teaching is the instructor's double bind--being with and without power. Within the classroom, teachers may seem all powerful. They may coerce students into doing what they would rather not by threatening bad grades which would harm their life chances (e.g., graduate school admissions and jobs). Yet in actuality, teachers cannot control many of the conditions of teaching critical for a positive outcome. Usually they cannot select the classroom or schedule, the size and composition of the class, the budget for extra resources, and so forth. The problem lies in the fact that students have high expectations of a seemingly powerful person who in actuality must frequently deal with scare resources. Indeed, teachers themselves are scarce resources. They must divide their time between the demands and needs of students on the one hand, their careers on the other. For in most universities, teaching, unlike research and politicking, goes unrewarded.

Teachers, then, are faced with an irresolvable conflict: they must in some manner convey to the students that they are approachable and effective as teachers. But they must also explain that they cannot always be available. The strain of this performance, combining personal attention and impersonal standards, sincerity and insincerity, drains the teacher even before treating the actual subject matter.

The treatment of the subject matter adds to this dilemma. As a scholar, the teacher is generally confronted with non-scholars. To adapt sociology to students presents an ongoing challenge because students' interests are different and constantly changing. Yesterday's approach is outmoded. To invent a new approach or to adapt one to the classroom, the professor must divert energy from scholarship and other more highly rewarded activities. The bind is apparent during pre-tenure years and it persists after tenure, since prestige among colleagues and financial rewards are usually divorced from the classroom. Within these realities, the classroom drama may vary, but it is always constrained. It is in fact that very professor, who has failed to gain recognition from peers, who directs energy toward the students as the one audience which may applaud. Conversely, of course, those academics who have been slighted by colleagues and administration may direct their frustration or hostility toward safer targets--the students.

One cannot properly observe a classroom without being aware of the many off-stage influences on the actors. Just as the teacher enters the classroom encumbered, so too do the students bring with them tensions of great complexity. They look for self-definition and validation, and simultaneously sense a frightening academic world often not of their choosing. An element of paranoia sets in, and often properly so as grades may determine their fates. They sit in the front row so that they may be remembered when medical school references are needed, or they sit in the back lest they be called on to speak of something they find obscure or trivial or seemingly irrelevant to their life purposes. What these purposes are is not usually clear to the students but they are conscious that the professor may be trying to change them in some way. This is perceived as a threat and consequently resisted. At a time when non-academic needs seem of paramount importance, difficulty in the classroom exacerbates students' personal troubles. This is especially true of sociology which students seem to expect will make their lives more meaningful. Few professors make explicit the connection between public issues and private lives, and so the abstract nature of sociology angers and disappoints the students.

Into the classroom, then, the professor and the students bring a host of selves. Here, to complicate matters, the structure of the classroom requires that they be in close proximity with others not of their own choosing. It is indeed remarkable that students can sit in a class for half a year without ever becoming acquainted with their neighbors. However, when it is recognized that the students are so arranged in rows that they can scarcely see each other, sensing the group but focusing straight ahead on the professor, it should not be surprising that tensions abound. The professor senses this also, too often attempting to resolve the difficulty by selecting one focal point, either in the form of an individual student or his own fantasy of an audience, ignoring as he does so the potentially dynamic forces in front of him.

Students' mixed motives for being in the classroom are further complicated by the person of the professor. The age and sex of the professor are important variables in determining students' response to him. Personality and timing are also important. At the same time as the group is exerting its influence on the individual student there is the dyadic drama between professor and student to account for. Both need consolation and support: teachers need to know that they are reaching their audiences. Students need to be recognized and encouraged for having learned something while at the same time resting assured that they have not lost the good will of their peers. Because gaining the approval of the professor oftentimes means losing the support of the group, coalitions can and do form, directed against the chief power holder, the professor. His perception of himself may not coincide with the perception of the class, and of even greater importance is his interpretation of the needs of his students which may not coincide with their needs at all.

Professors rarely have the opportunity to watch class dynamics from the outside. Neither is it traditional for professors and students to talk in settings other than the classroom. The consequences are immediate: it is much as if the professor were conducting a performance without fully understanding his own cues or recognizing student responses for what they are. Hence one can expect much dissonance. The interaction between professor and students is always complicated and tenuous, with very little required to permanently fracture the relationship.

It is not comforting to examine the classroom for, in microcosm, it mirrors the unmet needs, the unrealistic expectations, and the daily tragedies which characterize the social structure within which it is located. Each day students and professors pretend that learning experiences take place in a vacuum, but the reality is otherwise. The classroom is not sealed from the conflicts of the larger social context; indeed nowhere else is such conflict to be so immediately symbolized and perceived. While many professors of good will seek to make their classrooms sites of meaningful learning experiences, they are faced with great difficulty as it is often the case that the universities, like the society around them, seek more tangible profits than the nourishment of the life of the mind.

Thus, in the classroom a drama is staged with insufficient props and ambiguous scripts. Yet somehow the drama seems well worth playing (the Detroit workshop was evidence of this conviction). To play our parts better, we need to ask questions whose answers may well embarass many of us. Therefore, rather than stumbling through our roles, pretending we are masters of the performance, we need to accept that, like our students, we are amateurs, although, sometimes gifted ones. Such is the teaching-learning reality we all confront in our classrooms. Given this reality, how can we use the tools of our own craft to resolve some of the more pernicious aspects of the teaching dilemma? To this end, we turn now to a review of the literature of the classrooms.

Review of Literature

The central question guiding our inquiry asks: "How do we best conceptualize the classroom to understand why students learn or fail to learn?" A review of the literature reveals much research, largely on primary and secondary classrooms, that sensitizes us to some of the variables that might be incorporated into a model of classroom behavior and learning. Most of this research seeks to test the effects of student and faculty attributes on learning; or the effects of the classroom structure itself on learning. Although some researchers do attempt to trace the relationship between the functioning of classrooms and broader contextual and institutional arrangements (Boocock, 1972; Jackson, 1968; Waller, 1932; Hollingshead, 1949; Holt, 1967), their work represents a small proportion of published research on the classroom. The typical focus of the research is on social psychological processes within the classroom.

In their book Group Processes in the Classroom (1975), Schmuck and Schmuck approach the classroom from the point of view of the well established literature on small groups. Although the authors choose to focus their efforts exclusively on group processes that presumably occur in all classrooms, they do acknowledge that a complete understanding of the classroom requires that it be located within the context of broader social forces (Schmuck and Schmuck, 1975: 4):

> Each member of the classroom social configuration brings to it a special, unique set of characteristics, and since the classroom is only one part of the lives of its members, it is vulnerable to the influences of social forces surrounding it. The classroom group is directly and indirectly influenced by the total organization of the school building, the characteristics of the district in which the building lies, and the dynamics of the community in which the district is located. The different sociological environments of social class, race, and rural-urban differences create variances in the classroom's constitution.

As reflected in the Schmucks' work, the traditional social psychological literature on small groups emphasizes such topics as leadership styles in the classroom (Lewin, et al., 1939; White and Lippit, 1960; 1962); teacher expectation effects (Rosenthal and Jacobson, 1968; Brophy and Good, 1974); communications structures in the classroom (Bavelas, 1962; Leavitt, 1958; Adams, 1969); interpersonal attraction (Schmuck, 1966; Glidewell, et al., 1966; Lippit and Gold, 1959); and the emergence of classroom norms (Jackson, 1960). Although we appreciate the apparent generalizability of the small group processes discovered in these works, we consider the choice to view the classroom apart from still larger social processes and structures a conceptual limitation.

Another frequent concern of current literature on the classroom seeks to relate teacher and student personality attributes to learning outcomes (McKeachie and Lin, 1971; Mason, 1970; Bowers, 1962; Rees, 1969; Feenberg, 1972). Several of the studies with this emphasis suggest the hypothesis that there are distinctive student learning styles and teaching styles and that maximum learning takes place when there is the proper fit between the two (Mann, et al., 1970; Ringwald, et al., 1971; McKeachie, 1971; Grasha, 1972; Riechmann and Grasha, 1974). One line of thought here postulates that a transformation from teacher-centered to student-centered classrooms would advance the teaching-learning process (Silver, 1973). On the other hand, some would stress the interaction between teacher performance and student attributes. See the discussion of aptitude-treatment interaction by Solomon in Chapter IX.

Several other studies seek to construct classroom typologies by employing factor analytic techniques on questionnaire collected data. Solomon and Kendall (1976), for example, identify six classroom types:

(1) permissive and uncontrolled
(2) hostile, arbitrary, and regimented
(3) controlled, disciplined, academically oriented and supportive
(4) warm, oriented toward student expressiveness and creativity
(5) academically oriented, individualized interaction
(6) controlled, non-individualized, impersonal

Bennett (1976) likewise clustered classrooms on the basis of teacher questionnaire responses. This effort yielded twelve clusters differentiated along a continuum of teacher control and student autonomy. Similarly, Moos (1978) reports five clusters of high and junior high school classes: control oriented, innovation oriented, affiliation oriented, task oriented, and competition oriented. We might expect similar types to emerge were the same techniques used at the college level.

In the main, studies conducted by sociologists focus on aspects of the classroom's social structure. Accordingly, these studies shift the research concentration away from the personality characteristics of individuals. With respect to classroom interaction, for example, Adams and Biddle (1970) demonstrate that 63 percent of all student "emitters" in the sixteen classrooms they videotaped were located in three seats, one behind the other down the center of the room. If this center area is extended to the seats on either side of these three, nearly all student emitters are accounted for. This study replicates Sommer's (1967, 1968) earlier research on classroom ecology.

Other structural features of the classroom examined by sociologists include: bases of both formal and informal classroom *status* (Berger, et al., 1966; Gold, 1958; Katz, 1973; Glidewell, et al., 1966); *reward structures* (Coleman, 1962); the *teacher's role* (Waller, 1932; Dreeben, 1970); and *social control systems* (Smith, 1968; Leacock, 1969).

Few of the studies mentioned above appear to be guided by one or another specific theoretical viewpoint. An exception here is Karp and Yoels' (1976) observational study of selected college classrooms at a northeastern university. This study, written from a symbolic interaction perspective, explored the processes through which students and teachers formulate definitions of the classroom as a social space. One of their central arguments holds that certain organizational features of the classroom promote passivity or withdrawal on the part of some students. They name three as follows: (1) the "consolidation of responsibility"--students know that a small percentage of the class can be relied on to do most of the talking; (2) instructors are very unlikely to call on specific students; and, (3) students are tested infrequently. Given these characteristics of college classrooms, they conclude by speculating on the dilemma faced by college instructors who wish to have extensive classroom dialogues with students. Borrowing from Goffman, they suggest that college instructors define the classroom as an instance of focused interaction while students define it as an unfocused gathering.

Probably the most controversial sociological studies relate life in the classroom at all levels to the cultural, political, and economic organization of society. Recently for example, Bowles and Gintis (1976) have argued that a central function of the educational system in America is to sustain, reaffirm, and recreate the type of class structure required by a capitalist society. Their plainly Marxist thesis leads to an examination of the manner in which the day-to-day operation of classrooms--e.g., who gets to be in them (especially at the college level), which material is covered and how--reflects class interests. Other well-known studies (Jencks, 1972; Colemen, 1966) focusing on the relationship between persons' opportunities for social mobility and the structure of the educational system at all levels have much influenced recent social policy decisions.

The foregoing review of the literature, selective as it has been (we have excluded mention of the very large number of studies reporting on one or another pedagogical device), represents a fair sampling of the types of knowledge presently available about the structure and organization of the classroom. The need for more theoretically informed studies of the classroom, and especially the college classroom, seems plain. Beyond this, however, our literature review sensitizes us to the fact that what occurs in classrooms is part of a larger process; that the classroom is embedded in a larger social system, and should not therefore, be viewed in isolation.

Our literature review has served another heuristic function. We conclude from the studies discussed that a parsimonious model of how the classroom works must incorporate and describe the relations among the following factors:

(1) the sociocultural context (i.e., societally held values, attitudes toward education)
(2) the organizational context (i.e., the values and attitudes toward learning held in a particular school)
(3) the attitudes, styles, and personal characteristics teachers carry into the classroom
(4) the attitudes, styles, and personal characteristics students carry into the classroom
(5) the structural characteristics of the classroom itself.

In the following sections we will elaborate a model based on these variables, argue its significance for understanding different patterns of learning outcomes, and specify important areas for empirical research revealed by the model.

Theoretical Assumptions and Framework

It is quite clear that paradigmatic orthodoxy is not a useful way of selecting a guiding framework for our task. We are assuming a dialectic between structure and process; we are assuming a systematic interdependence between component parts of the learning environment; we are assuming conflict and power relationships in the classroom which give rise to much of the drama already explicated; we are also assuming consensus and shared definitions of the situation under some conditions; and we conceptualize the classroom not as an entity but rather as a process in which evaluation, definition, and redefinition of the situation occur, becoming recurrently salient in the process of interaction. In short, we need an elaborate kit of theoretical constructs to move us toward greater clarity in analyzing this problem.

The following is our schematic depiction of the teaching-learning process which we will use to map the classroom in its social context, while ultimately trying to explain differential patterns of learning outcomes as our dependent variable.

Elaboration of the Scheme

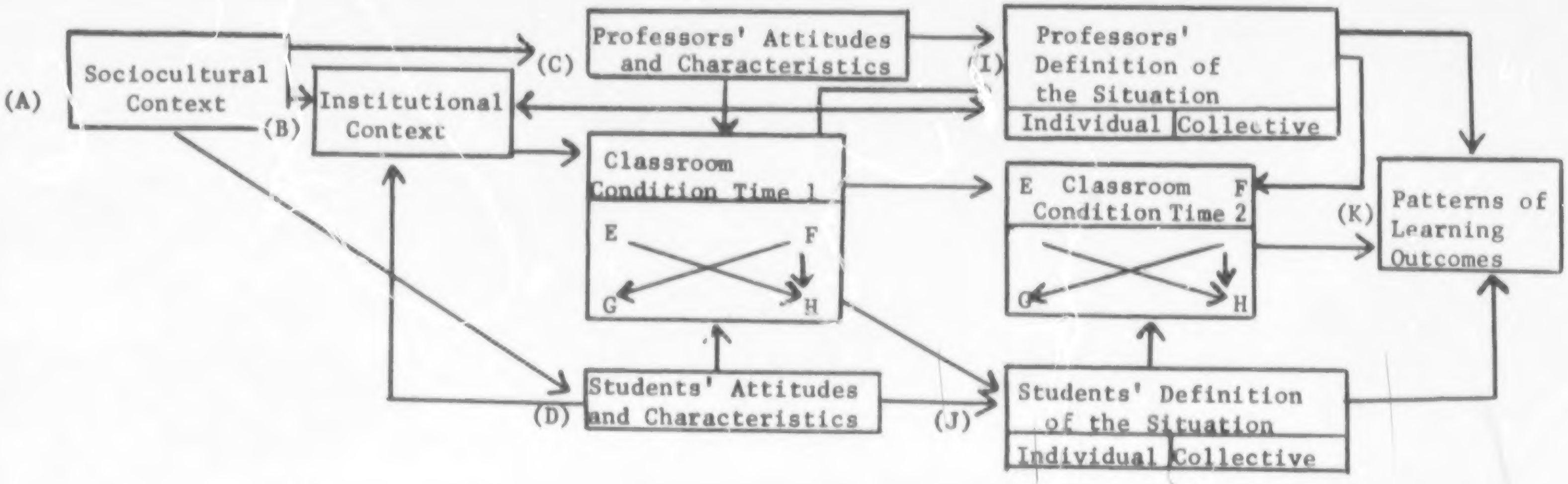
Our scheme, a process model, employs several contextual variables which have both direct and indirect effects on learning outcomes which cannot be predicted for a given classroom. Each of these contextual variables interacts in complex fashion to bring structure and process face to face. Our aim in this section is to demonstrate some of this dialectical complexity. We will begin by briefly identifying each of the variables and then characterizing them in terms of their component parts.

A. Sociocultural Context:

By this we are referring to the complex of values, norms, and general cultural traditions which the educational institutions inherit and perpetuate in the historical process. For example, the fact that America has a legacy of pragmatism and vocationalism expresses itself in centers of higher learning through a structuring of the kinds of educational options likely to emerge. Any listing of component parts would include value complexes, norms, beliefs, symbols, sentiments, and attitudes prevailing in a given historical epoch.

B. Institutional Context:

This refers to the nature of educational functions served by an organization--both manifest and latent· e.g., liberal arts, engineering, or technical schools. Important here also is the location of the particular educational organization in the evaluation scheme operating for its category. The context is composed of not only tangible resources but also concomitant psychological orientations. The following factors operate as components of the institutional context: (1) age and prestige, (2) credentials of the faculty, (3) admissions standards, (4) region of the country, (5) intellectual ethos, (6) faculty remuneration and morale, (7) departmental structure (includes sex ratio, Chair or Headship, percent tenured, age structure, reputation of the department in the school's history), (8) teaching resources, (9) size of classes, (10) degree of interaction in student body between racial, sexual and age groups, (11) size of the school, (12) reward structure for faculty (teaching vs. research), (13) administrative structure, (14) teaching load, (15) committee load, (16) advising load, (17) town/gown


(A)
Sociocultural Context
(B)
Institutional Context
(C)
Professors' Attitudes and Characteristics
(I)
Professors' Definition of the Situation
Individual
Collective
Classroom Condition Time 1
E
F
G
H
E
Classroom Condition Time 2
F
G
H
(K)
Patterns of Learning Outcomes
(D)
Students' Attitudes and Characteristics
(J)
Students' Definition of the Situation
Individual
Collective

A. Sociocultural Context

B. Institutional Context

C. Professors' Attitudes and Characteristics

D. Students' Attitudes and Characteristics

E. Physical Environment

F. Demographic Factors

G. Course Content

H. Interaction Process Variables

I. Professors' Definition of the Situation

J. Students' Definition of the Situation

Figure 1: SCHEMATIC MODEL OF CLASSROOM AS SOCIAL PROCESS (1)

relations, (18) library quality, (19) Greek/nonGreek residences, (20) faculty interaction patterns (formal/informal).

C. Professors' Attitudes and Characteristics:

Most of the literature on the classroom has fixed on this variable as though it were the only one operative. Again, placing even this apparently simple trait variable in its sociological context reveals at least two sets of factors: attitudinal orientations of the instructor and his social profile.

Attitudinal Characteristics:

All of the following will enter into the classroom interactions under particular conditions: (1) degree of professional socialization, (2) professional self concept, (3) perception of goodness-of-fit with prestige of organization, (4) esteem for and from colleagues, (5) degree of commitment to teaching, (6) cosmopolitan-local orientation, (7) degree of integration in organization--social and academic, (8) level of satisfaction with remuneration, (9) theoretical orientation, and (10) degree of identification with institutional goals.

Social Characteristics:

The following are objective status and role attributes of significance in the classroom: (1) race, (2) sex, (3) ethnicity, (4) social class background, (5) age, (6) religion, (7) tenure status, (8) marital status, (9) recency of training, (10) teaching style, (11) speech and movement pattern, (12) reputation on campus, (13) physical health, (14) physical size.

D. Students' Attitudes and Characteristics:

One of the most powerful structures confronting the classroom instructor is that of the fabled student body. In analyzing key features of our ultimate raison d'etre, we developed the following inventory which, as with professors, can be categorized as attitudinal (psychic) or social characteristics.

Social Profile:

(1) social class origins, (2) race, (3) sex, (4) religion, (5) quality of high school preparation, (6) age, (7) employment status, (8) commuter status, (9) marital status, (10) measured I.Q., (11) overall GPA, (12) social mobility experience, (13) physical health, (14) life styles, (15) subcultural membership, and (16) cognitive styles.

Attitudinal Features:

(1) level of motivation to learn, (2) degree of intellectual socialization, (3) attitude toward subject, (4) career aspirations, (5) level of life satisfaction, (6) self concept, (7) satisfaction with parental relationships, and (8) attitude toward authority.

We turn now to "Classroom Condition--Time 1" which is clearly the locus of the action. This is the critical social space--one which brings together all the values, attitudes, characteristics and complexes already outlined. In general, this space may be characterized as comprising a physical environment, a human aggregate with a demographic profile, course content and group process factors. We will briefly outline key features of each.

E. Physical Environment:

As sociologists, we are not inclined to give much credibility to the notion that the physical structures of our classrooms are a powerful influence on the quality of the social dynamic. Yet, how many times do we get complaints about just such factors? And how many times do we marvel at the range of quality we observe in physical facilities as we visit in different schools? We are proposing here that physical space is a key feature in terms of at least the following: (1) ventilation, (2) space/people ratio, (3) shape, (4) flexibility of seating and walls, (6) seating comfort, (7) technical resources, (8) pollution level (smoking/non-smoking), (9) noise level, (10) cleanliness and decor.

Sensitivity to the impact of these factors on intellectual productions can suggest intervention which might otherwise be deemed trivial. Anything which potentially diminishes the quality of the classroom interaction is not trivial and should be rectified wherever feasible.

F. Demographic Factors:

(1) ratio of majors/nonmajors, (2) sex ratio, (3) number of credits, (4) whether required or elective, (5) length of class, (6) time of day, (7) time of year, (8) ratio of freshmen to upperclasspersons, (9) size of class, (10) racial homogeneity, and (11) class homogeneity.

G. Course Content:

(1) work load, (2) level of abstractness, (3) level of complexity, (4) perceived practicality, (5) directness of application, (6) formality of syllabi (expectations), (7) ratio of field work/classroom, (8) grading norms, (9) prerequisite/cumulative, (10) sequence in program, and (11) social inquiry or information transmission motif.

H. Group Process Variables:

Thus far we have identified the kinds of dimensions and variables which, from a review of the literature and our own reflection, seem important. At this point, we need to specify more exactly how those elements interact with each other. This brings us to the actual dynamics of the group, characterized by: (1) communication patterns, (2) leadership and influence, (3) attraction and friendship networks, (4) group cohesion, (5) emergent norms, (6) teacher expectations, (7) task and maintenance performances, (8) emergent consolidation of responsibility, and (9) teaching/ learning styles.

I. Definitions of the Situation:

Given that every classroom will have this mix of physical, institutional, attitudinal, and perceptual variables, participants will arrive at a definition of the situation. Such definitions are, of course, being made and revised continuously. Students make their own private assessments of what the class is all about. They also talk with each other and through peer networks arrive at a collective definition of the situation.

Professors engage in the same process, though from drastically different perspectives. They too make private judgements about each class and construct a definition of the situation. In conversations with colleagues, professors arrive with expectations, motivations, and the contingencies already discussed--all of which influence the constructed reality of the class room.

It is on the basis of these definitions that students and professor fashion appropriate responses. Professors may decide to lecture primarily, or to have discussion groups. Students may decide to sleep in class, boycott every third lecture, dominate the discussion, etc. These responses are mostly experimental at the beginning of the semester, but at some point relatively early in the process, there may be a transformation in classroom behavior. The mood, tenor, level of trust, level of performance, degree of responsibility--all will be a function of the earlier definition of the situation.

In turn, the professor will be defining and redefining based on feedback experiences in the initial encounters. The degree to which major revisions are made will usually be a function of commitment to initial definitions, resistance from students, amount of energy available for the task, etc.

This process of negotiation will function most importantly at the beginning of the interaction. Shortly into the course, the teaching-learning environment of the classroom will crystallize. That is, the definitions of the situation of both students and teachers will largely be determined--expectations will harden, patterns will routinize, and classroom norms will take on an external facticity.

J. Pattern of Learning Outcomes:

Up to this point we have been analyzing the teaching-learning process of the classroom, dissecting a process having affective and cognitive consequences for the participants, both students and professors. Those of us who teach will readily acknowledge that each classroom is unique and produces a different pattern of learning. Some produce excited, romantic enthusiasts for the subject matter. Others produce hostility and terminal boredom. Yet others produce technical competence but little enthusiasm which might last beyond final examinations. Ultimately, we would all ideally like to produce competent, enthusiastic, intellectually curious students who cannot wait to take another of our courses or who hasten to apply their sociological insights to every conscious--and even subconscious--interaction.

Given this ideal and, perhaps more importantly, the gap between actual and ideal results, how can we use this model to diagnose teaching/learning environments and enhance the probability of achieving ideal patterns of learning in our students? It is to this question that we now turn.

Application of the Model

This theoretical schema can be applied to the teaching craft for two very different purposes. The first application is that of sensitizing the instructor to the complexities of the classroom and helping to diagnose the variables which are amenable to immediate, teacher-determined intervention. For example, while class size and physical space may be bureaucratically imposed, interaction patterns in the classroom are not. Nor are stuffy, smoke-filled rooms imposed, any more than ambiguous teacher expectations. The instructor can take this model with its list of variables, consider each one in turn, rank order them for importance in the particular case and arrive at a decision as to which ones can be improved in the short-term (such as smoking), which are middle-range (such as a more flexible room), and which are long-term (such as class size and quality of students).

A second use of the model is that of generating a research agenda for the classroom such that we may follow the urging of E. K. Wilson (1977) to "banish the conjunction" between research and teaching. The model now can be used to move us from being intuitive amateurs to informed professionals. The research task is waiting under the banner: "knowledge needed."

Knowledge Needed: A Sample of Researchable Hypotheses

It is distressing to discover how much we take for granted in our classrooms. Even the most honored verities--e.g., authoritarian classrooms produce poor learning outcomes--do not have sound research support. What we are left with is an almost inexhaustible research agenda relating to classroom interactions. With the number of variables we have sketched, we can readily generate a sample list. However, the relative importance of hypotheses or sets of hypotheses will probably be a function of specialized circumstances and orientations of instructors. The following are designed only as examples of theoretical linkages which need empirical testing.

1. The greater the discrepancy between the institutional reward structure and the instructor's expectations, the lower the level of teaching commitment.

2. There is a negative association between the level of professional commitment to teaching and the size of the employing institution.

3. The greater the academic prestige of the institution, the lower the level of commitment to teaching if size is held constant.

4. The greater the heterogeneity of student educational background, the lower the Quality of Learning Outcomes (Q.L.O.).

5. The higher the concentration of Greek fraternal organizations, the lower the intellectual motivation of the student body.

6. The greater the attitudinal and value homogeneity of students in a classroom, the more quickly and unanimously will they arrive at a collective definition of the situation.

7. The greater the discrepancy between the attitudes and characteristics of professors and students, the slower the emergence of the definition of the situation and the higher the Q.L.O.

8. The greater the clarity and rigor of teacher expectations, the faster the emergence of the definition of the situation and the higher the Q.L.O.

9. The higher the level of negotiated sharing of responsibility for the classroom productions, the higher the Q.L.O.

10. There is a negative association between the number of classroom stars and the Q.L.O.

Summary and Conclusion

We began this chapter with the inquiry: "why is it that some classroom contexts produce high levels of learning while others do not?" Our own teaching experience and a review of the literature did not reveal much that could be called reliable knowledge on the sociology of the classroom, except of the most general kind. Our thinking about teaching-learning environments led us to depict them as dramatic stages on which teaching and student roles are joined, sometimes tragically. It also led us to a deeper appreciation of the complexity of teaching and learning.

Our review of the literature exposed a disturbing dearth of sociological studies of classrooms which might have shifted the analysis from its dominant psychological and social psychological orientation to its larger socio-cultural context. Those studies which have spanned the different levels of analysis--cultural, social, political-economic and psychological--have focused more on wholistic educational institutions than on their component parts, such as classrooms.

This brought up the need for a theoretical framework which would link the varying levels of analysis and serve as both a diagnostic tool for instructors and administrators, as well as indicating the gaps in our empirical knowledge. We developed a parsimonious model which accommodates both structural and process variables. "Pattern of Learning Outcomes" is our dependent variable, in keeping with our basic statement of the problem. We elaborated the model with an inventory of component variables and suggested at least two applications of the model, not only for those who profess sociology but for all who teach as a profession. We have not tried to operationalize any of the variables introduced here. We assume that our colleagues are competent methodologists capable of taking up the research task where interest dictates.

In undertaking this task, we gained a new appreciation of Sorokin's classic introduction to Social Mobility (1927: 414):

> The task before me is not unlike that of a physiologist who, though he has studied but a few small spots in the circulatory system, must yet on this basis construct the whole process of blood circulation of the organism. Such an enterprise may naturally lead to mistakes. And yet, since the necessary material is absent, somebody must venture to construct a theory in order to stimulate further studies.

He went on to say that it was "high time to abandon speculation for the somewhat saner method of collecting facts and studying them patiently." Sorokin's comments, although referring to processes of social stratification, apply equally to a sociology of the classroom.

For those of us who profess sociology, it is ironic that the least researched topic is our own workplace, the classroom. The research agenda is clear; the knowledge awaits the knower.

References

ADAMS, R.S. (1969) "Location as a Feature of Instructional Interaction." Merrill-Palmer Quarterly of Behavior and Development 15:309-321.

ADAMS, R.S. and B.J. BIDDLE (1970) Realities of Teaching: Explorations with Videotape. New York: Holt, Rinehart and Winston.

BAVALES, A. (1962) "Communication Patterns in Task Oriented Groups." pp. 669-682 in D. Cartwright and A. Zander (eds.) Group Dynamics. Evanston, IL: Row, Peterson and Co.

BENNETT, N. (1976) Teaching Styles and Pupil Progress. Cambridge, MA: Harvard University Press.

BERGER, J., et al. (1976) "Status Characteristics and Expectation States." pp. 47-73 in Berger, Zelditch and Anderson (eds.) Sociological Theory in Progress. Boston: Houghton-Mifflin.

BOOCOCK, S.S. (1972) An Introduction to the Sociology of Learning. Boston: Houghton-Mifflin.

BOWERS, M.D. and R.S. Soar (1962) "The Influence of Teaching Personality on Classroom Interaction." Journal of Experimental Education 30:309-311.

BOWLES, S. and H. GINTIS (1976) Schooling in Capitalist America: Education and the Contradictions of Economic Life. New York: Basic Books.

BROPHY, J. and T. GOOD (1974) Teacher-Student Relationships: Causes and Consequences. New York: Holt, Rinehart and Winston.

COLEMAN, J. (1962) "Reward Structure and Allocation of Effort." pp. 119-132 in J.P. Criswell, et al. (eds.) Mathematical Methods in Small Groups Processes. Stanford, CA: Stanford University Press.

COLEMAN, J., et al. (1966) Equality of Educational Opportunity. Washington, D.C.: U.S. Department of Health, Education and Welfare; Office of Education.

DREEBEN, R. (1970) The Nature of Teaching. Glenview, IL: Scott, Foresman.

FEENBERG, L. (1972) "Faculty-Student Interaction: How Students Differ." Journal of College Student Personnel 13:24-27.

GLIDEWELL, J.D., et al. (1966) "Classroom Socialization and Social Structure." pp. 221-257 in M. Hoffman and L. Hoffman (eds.) Review of Child Development Research. New York: Russell Sage Foundation.

GOLD, M. (1958) "Power in the Classroom." Sociometry 25:50-60.

GRASHA, A. F. (1972) "Observations on Relating Teaching Goals to Student Response Styles and Classroom Methods." The American Psychologist 27:144-147.

HOLLINGSHEAD, A. (1949) Elmtown's Youth. New York: John Wiley.

HOLT, J. (1967) How Children Learn. New York: Pitman.

JACKSON, P. (1968) Life in Classrooms. New York: Holt, Rinehart and Winston.

JENCKS, C. (1972) Inequality: A Reassessment of the Effects of Family and Schooling in America. New York: Basic Books.

KARP, D.A. and W.C. YOELS (1976) "The College Classroom: Some Observations on the Meaning of Student Participation." Sociology and Social Research 60:421-439.

LEACOCK, E. (1969) Teaching and Learning in City Schools. New York: Basic Books.

LEAVITT, H.J. (1958) "Some Effects of Certain Communication Patterns on Group Performance." pp. 546-563 in E.E. Maccoby, et al. (eds.) Readings in Social Psychology. New York: Holt, Rinehart and Winston.

LEWIN, K., LIPPITT, R. and R.K. WHITE (1939) "Patterns of Aggressive Behavior in Experimentally Created 'Social Climates'." Journal of Social Psychology 10:271-300.

McKeachie, W.J. and Y. Linn (1971) "Sex Differences in Student Response to College Teachers: Teacher Warmth and Teacher Sex." American Educational Research Journal 8: 221-226.

MANN, R.D., et al. (1970) The College Classroom: Conflict, Change and Learning. New York: John Wiley.

MASON, J.L. (1970) "Study of the Relationship Between the Behavioral Styles of Classroom Teachers and the Quality of Teacher-Student Interpersonal Relations." Educational Leadership 28:49-56.

MOOS, R.H. (1978) "A Typology of Junior High and High School Classrooms." American Educational Research Journal 15:53-66.

REES, R.D. (1969) "Dimensions of Students' Points of View in Rating College Teachers." Journal of Educational Psychology 60:476-482.

RIECHMANN, S.W. and A.F. GRASHA (1974) "A Rational Approach to Developing and Assessing The Construct Validity of a Student Learning Styles Scale Instrument." The Journal of Psychology 87:213-223.

RINGWALD, B., R. MANN AND W. J. McKEACHIE (1971) "Conflict and Style in the College Classroom: An Intimate Study." Psychology Today (February).

ROSENTHAL, R. and L. JACOBSON (1966) "Teachers' Expectancies: Determinants of Pupils' IQ Gains." Psychological Reports 19:115-118.

SCHMUCK, R. (1966) "Some Aspects of Classroom Social Climate." Psychology in the Schools 3:59-65.

SILVER, R.J. (1973) "Letting Go in the Classroom." Sociological Inquiry 43:169-178.

SMITH, L. and W. Geoffrey (1968) The Complexities of an Urban Classroom. New York: Holt, Rinehart and Winston.

SOLOMON, D. and A.J. KENDALL (1976) Individual Characteristics and Children's Performance in Varied Educational Settings: Final Report. Spencer Foundation Project, Montgomery County (Maryland) Public Schools. Rockville, MD.

SOMMER, R. (1967) "Classroom Ecology." Journal of Applied Behavioral Science 3:489-508.

---(1969) Personal Space. Englewood Cliffs, N.J.: Prentice-Hall.

WALLER, W. (1932) The Sociology of Teaching. New York: John Wiley.

WHITE, R. and R. LIPPITT (1960) Autocracy and Democracy. New York: Harper and Row.

---(1962) "Leadership Behavior and Member Reactions in Three 'Social Climates.'" pp. 527-553 in D. Cartwright and A. Zander (eds.) Group Dynamics. Evanston, IL: Row, Peterson.

WILSON, E.K. (1977) "Sociology: Scholarly Discipline or Profession?" Keynote Address introducing the meetings of the ASA Section on Undergraduate Education; Annual Meeting of the American Sociological Association. Chicago (September 8).

To: FUDGE (Foundation for Undergraduate Development and General Education)

From: James A. Davis (Harvard University), Richard Dukes (University of Colorado), and William A. Gamson (University of Michigan)

Re: R & D Proposals for Assessing Nontraditional (Interactive) Modes of Sociology Instruction

Gentlepersons:

We were delighted to learn that FUDGE has chosen to start a major new program of research and development on the teaching of the social sciences. We think you are wise to spend some time analyzing what your research priorities should be, and we appreciate the opportunity to offer our advice. We will confine our remarks to the broad sub-area which you have designated as "Learning Beyond the Classroom."

Our concern is with a series of teaching and learning modes that frequently involve experiences both in and out of the classroom. We take the interactive nature of the mode rather than its setting as the defining characteristic. By an interactive mode, we mean one in which the student is actively involved in performing some task--be it playing a game, interacting with a computer, participating in a field experience, collecting original data, or other analogous activity. In all of these, we assume that the students will at various points interact with instructors and peers to discuss, analyze, and write about their activities.

For convenience, we suggest that you divide these activities into three broad sub-classes as follows:

1. Computer explorations. We have in mind here those learning exercises in which students interact with computers. For example, a class might have available a prepared data set and a series of questions that can potentially be answered with such a set. Individuals then have access to a computer terminal to work on the exercise individually or in small teams, later convening to share and compare their analyses and ultimately to write them up. (For a fuller description of an exercise of this sort, see James A. Davis, "Teaching Social Facts with Computers" in *Teaching Sociology*, April 1978.)

2. Games and simulations. Here we include non-simulation games such as those of Layman Allen (*Equations*, *Wfn' Proof* or *Theories and Queries*); simulation games; role playing exercises such as mock political conventions

or UN sessions. We also include here any other classroom techniques that are "keyed" to use Goffman's (1974) concept.

3. Practicum experiences. We include here those out-of-classroom experiences that are explicitly defined as having an educational or training goal. We emphasize the explicitness because we would not deny that one may learn from virtually any experience regardless of whether it has a formal learning objective. The presence of a formal learning objective is necessary to put some limits on what is relevant here. A reasonable operational definition is that the learning experience have a product of some sort attached to it that is assessed by peers or instructors. Such things as research practicums, field work placements, and internships would typically fall into this category.

Suggested Program Priorities

The experience of teachers who devise interactive scripts is a good basis for identifying what research and development priorities should be. We suggest that you worry about three central issues of concern to such practitioners: evaluation, costs, and transcending the immediate experience.

1. The assessment of "seeing connections." The users of these techniques frequently indicate broad learning objectives that are difficult to assess with conventional testing instruments. They may mention such things as "getting students to see patterns," "learning how to think like a social scientist," "seeing the forest as well as the trees," "getting the idea of a social system," "gaining intuitive grasp," or "learning the music as well as the words."

Bloom, in his well known taxonomy of educational objectives uses the term synthesis to characterize these objectives:

> The putting together of elements and parts so as to form a whole. This involves the process of working with pieces, parts, elements, etc., and arranging and combining them in such a way as to constitute a pattern or structure not clearly there before.

We have no intention of suggesting that such a learning objective is unique to the interactive learning techniques we are considering; clearly, such objectives are likely to be present in many conventional teaching situations as well. We emphasize this objective here because it is so central to the claims made for the special merit of interactive techniques in producing such learning. Raser (1969), for example, quotes Campbell's distinction between "distal" and "punctiform" knowledge in justifying the use of game simulation.

> We confront a collection of fragments--bits of punctiform data, each of which is uninterpretable--and suddenly we see the entire pattern or context. Common expressions used to describe this experience include "insight," "revelation," "seeing how it all fits together," or "having it suddenly all make sense." They all express the recognition that when an entire context or pattern is grasped, each part of the pattern is also more clearly apprehended.

Those who use computer exploration would argue, similarly, that a student who has carried out his own analysis of a data set will understand the connections among the variables in a more fundamental way than one who has simply read such an analysis by another. Similarly, practicum experiences are frequently intended to link theoretical ideas to concrete examples, thereby enabling a student to see the general in the particular and vice versa.

To sum these claims up in a phrase, the learning experience is designed to help students make connections. There are no instruments currently in circulation that people working in this field would generally accept as assessing such learning in a reliable fashion. We believe that FUDGE could make a major contribution by sponsoring studies that would develop and test a variety of such instruments.

A variety will surely be necessary because it is unlikely that context-free measures will prove viable; rather, this general learning objective will need to be operationalized differently for different contexts. On the other hand, there is some hope that when a variety of such measures have been developed and established as useful in given contexts, they will prove useful (perhaps with some modifications) in additional, similar contexts. The researcher who is hoping to improve the effectiveness of an interactive learning technique would then have available a series of possible tools for assessing the degree to which students' skills in making connections are improving, in the teaching situation of interest.

We suggest that the most successful methods of assessment of this learning outcome are likely to have the following properties:

1. An emphasis on using or applying one's knowledge or understanding rather than on recognizing or recalling something;
2. An emphasis on the analysis of a new situation not previously encountered rather than mastery of the original situation in which the connections were learned;

3. An emphasis on analyzing a complicated and noisy situation where confusion is likely to befall those who are unable to make the requisite connections.

For each mode of interactive learning, we can think of a number of possible means of assessment. After several exposures to computerized data analysis problems, the student should be better able to:

(a) write a term paper in another sociology course where quantitative data are available;

(b) provide a cogent critique of a newspaper or magazine article involving a fallacious statistical argument;

(c) draw a schematic causal diagram of a verbal argument on a new topic;

(d) suggest the logic for testing a proposition advanced by an instructor in a substantive course;

(e) spot the anomaly in a data set in which an error has been deliberately inserted.

After exposure to and analysis of a game simulation, a student should be better able to:

(a) analyze what would happen if a rule change were introduced in the game;

(b) analyze what would happen if a given change in social structure occurred in the real world process being simulated;

(c) analyze the strategic options available for actors in the process being simulated;

(d) develop one's own game to simulate some analogous social process.

After a field experience, a student should be better able to:

(a) offer advice to a person with a problem of the sort encountered by people in the field situation;

(b) provide an analysis of a case study which raises in a different setting, issues similar to those encountered in the field situation;

(c) write a field report on a new observation setting;

(d) write a critique of a published research report on the problem or area studied.

We offer these suggestions in the spirit of illustrative hypotheses. We would hope that as FUDGE supports research in this area, we will learn more about which suggestions are most useful in assessing increases in distal (synthesizing) knowledge.

2. Controlling the cost of interactive techniques. Teaching in this mode frequently is more costly than conventional teaching and some of the costs are hidden. First, there are preparation costs. Expensive assistance may be required to prepare data sets for use in computer exploration and terminals must be provided. Even a relatively small class can use quite a bit of computer time which must be charged to a Departmental or other account which invariably has limited funds. Game simulations may require the purchase of fairly expensive materials and typically require considerable time for the instructor to prepare. A great deal of advance planning and preparation is necessary to put in place the mechanisms for a successful field experience.

These techniques also generate hidden costs in time and energy of the faculty members who use them. They are, typically, more time consuming to use than conventional techniques. We suspect that one of the mechanisms is a changed student-faculty relationship which often occurs. Individual students develop relationships with faculty members that lead them to seek advice on a variety of academic, career, and even personal matters. The interactive mode of teaching, we suggest, spills over in this and in other ways, tending to increase the required commitment of faculty time.

The evidence for such a spill-over effect is largely anecdotal and we don't mean to suggest that such costs are inevitable or occur under all conditions. But once recognized as a potential problem, costs can be anticipated and ways found of reducing them. Systematic use of peer feedback can, at least potentially, relieve the instructor of some of the extra-burden of using such instructional techniques. We would hope that FUDGE would sponsor research that would explore different modes of peer instruction and their strengths and shortcomings in different situations. Such research could tell us what conditions must be met for various sorts of peer instruction to work effectively. Are such techniques effective in large classes, for example? Do they relieve the instructor or do they add still another difficult supervisory responsibility without releasing time from other tasks?

While costs are a serious problem in this style of learning, this very problem gives FUDGE important opportunities for research. Costs are very researchable, as Economics and Operations Research show. And when useful findings occur, they can be disseminated widely. Research on the most efficient ways to organize and implement these teaching techniques should have high priority. More bluntly, we feel the time has come to shift from asking whether to do it to asking how to do it more cheaply.

3. The problem of over-concreteness. We want students to be able to recognize something generic in a concrete example. In the case of a field experience, we want students to view their own personal experience in a larger context, to be able to see something more general about the particular. In game simulations, we want students to be able to perceive the relationships in the real world that the game is simulating.

It is clear from studies of game simulation, at least, that not all students are able to move from the concrete to the generic. In summarizing the many studies of the effectiveness of game simulations, Livingston and Stoll (1973) conclude that:

> Students of low academic ability or achievement usually do much better at learning to play a simulation game than at learning *from* the game; their greatest difficulty is in transferring their understanding of the game to the real situation that the game represents.

There are many researchable questions suggested by this general problem. Are there conditions under which engagement in a concrete task works against this learning objective, leading students to perseverate on idiosyncratic aspects of the concrete case? Davis describes cleaning and simplifying the data sets used for student computer analyses. In doing so, he is attempting to protect students from becoming so bogged down in marginal or non-existent relationships, that they will miss the central ones present in the data. Does such a method work, and is it part of a more general class of simplifying techniques that help to spotlight generic aspects of concrete material and, thereby, increase the chances of fruitful generalization?

Similarly, Coleman (1966) argues that game simulations are caricatures of social life--indeed, that this is one of their prime virtues. Do such caricatures, with their deliberate exaggerations of some features, facilitate generalization? How does one simplify a field experience to increase the likelihood of such generalization?

Ironically, one of the virtues of interactive learning techniques may turn out to be a handicap: their ability to produce student engagement.

We are accustomed to viewing engagement of students in a learning task as an unadulterated blessing but those who use these methods sometimes experience problems of over-engagement. Detachment may be necessary to enable the student to step back from an intense involvement to see events in a perspective that promotes generalization. Does the very engagement which these teaching methods often produce present an obstacle, under some circumstances, to the cognitive learning objectives they are designed to promote?

Areas of Research to De-emphasize

Our advice may be more helpful if we also indicate some areas of research that should be given very low priority. We see no need for further research to establish student preferences for these techniques. It is well established, at this point, that students tend to like them more than conventional classroom instruction. Writing of the seven year experience of the Academic Games Program at Johns Hopkins University, Livingston et al. (1973) conclude that "the most consistent finding of the Academic Games Program's research with games in the classroom is that students prefer games to other classroom activities." Seidner (1976), in her review of the state of the art, concludes that:

> Not only do students report that they prefer gaming activities, even when compared to other innovative teaching techniques, but behavioral indicators indicate that they mean what they say (lower absence and drop-out rates, increased participation).

While evidence for the motivating effects of field experiences and computer play is more anecdotal, it is quite consistent with the above conclusions on game simulations. We see little point in gathering more evidence on this point. The problematic aspects of the techniques concern cognitive outcomes and these are the ones FUDGE should focus on in supporting future research.

Equally low priority should be given to conventional evaluation studies that compare a control group using a conventional classroom with an experimental group using some interactive technique. We don't think there is much more to be learned from this research paradigm. We doubt that fruitful, context-free generalizations can be found demonstrating that one technique is uniformly better than another, even for specific learning objectives.

Our alternative approach accepts the fact that these techniques show no evidence of general inferiority to conventional techniques and that particular games, computer exercises, and field experiences will vary greatly among themselves in effectiveness. The focus should be on the conditions under which given interactive techniques are most and least appropriate. We need to know the contextual variables that maximize the effectiveness of a given method. More relevant than a comparison with a conventional classroom is a comparison between the use of a given technique in various contexts (required vs. non-required courses, variations in type of product, different systems of peer feedback, etc.).

In sum, we suggest that your research agenda should address questions that arise among those already committed to using such techniques in teaching sociology. We have tried to suggest what some of these questions are from our own perspective as practitioners and consumers of research in this field. The critical question is not whether these techniques are better or worse than something else but for what purposes and under what conditions they work most effectively.

We append a list of references for material cited in this memorandum.

References

BLOOM, B. (1956) A Taxonomy of Educational Objectives. New York: D. McKay Co.

COLEMAN, J. S. (1966) "In defense of games." American Behavioral Scientist (No. 2): 3-4.

DAVIS, J. A. (1978) "Teaching social facts with computers." Teaching Sociology 5: 235-258.

GOFFMAN, E. (1974) Frame Analysis. Cambridge, Mass: Harvard University Press.

GREENBLAT, C. S. (1973) "Teaching with simulation games: a review of claims and evidence." Teaching Sociology 1 (October): 62-83.

LIVINGSTON, S. A. et al. The Hopkins Games Program: Final Report on Seven Years of Research. Report No. 155, Center for the Social Organization of Schools. Baltimore, Md.: The Johns Hopkins University.

LIVINGSTON, S. A. and C. S. STOLL (1973) Simulation Games: An Introduction for the Social Studies Teacher. New York: Free Press.

RASER, J. R. (1969) Simulation and Society. Boston: Allyn and Bacon, Inc.

SEIDNER, C. J. (1976) "Teaching with simulations and games," pp. 217-251 in The Psychology of Teaching Methods. Chicago: The National Society for the Study of Education.

Grading and the Process of Developing Unfettered, Self-Governing Persons

Bill Ewens

Michigan State University

Schooling is never a neutral process. It can either serve to make persons more unfettered, self-governing, and capable of the practice of freedom or it can serve to domesticate them and make them more reliant upon outside authority.

Currently in this country, the use of student grades as a relevant measure of learning is a vital part of this latter domestication process. Authoritarian societies reproduce themselves by socializing persons to rely on outside authority as a basis for making decisions rather than relying upon their own natural feelings and impulses. To accomplish this, persons must be created who are guided by a rigid conscience and a compulsive morality and who routinely censor their own natural impulses in instances where these impulses are inconsistent with existing social rules. Moreover, it is necessary that these persons have a low opinion of themselves and their abilities to take responsibility for their own lives. They must, in Reich's words, "fear freedom" and be systematically trained to rely on experts to guide their lives. Spontaneity is to be avoided at all costs. Just as strong leadership and coercive social controls are necessary according to authoritarian ideology to keep the delicate fabric of society from being torn asunder, so too must individuals be eternally vigilant to repress the least sign of independence of rebelliousness in their own souls.

If authoritarian societies are to reproduce themselves successfully then schools must become a general microcosm simultaneously reflecting and reinforcing the general structure of the overall social order. Intrinsic satisfaction from learning and spontaneous problem posing by students become threats to the basic integrity of the system. Also to be abhorred are student demands for a voice in controlling their own education. Instead students must be trained to rely on teachers to pose the relevant problems and to direct them to the proper solutions. The social relations of both the classroom and the school must be hierarchical and controlled in a top down fashion. Grades, like wages in industry, become the currency of the system and the primary reinforcers of proper conduct. In preparation for other adult roles in an authoritarian society students learn to function in an alienating social environment where there is a basic split between what they must do to be successful students, on the one hand, and what they must do to become more fully functioning human beings, on the other hand.

Current Research on Grading

Most current educational research takes for granted the present system of schooling in the United States in which student grades are a basic characteristic of the teaching-learning process. Science is thus placed in a subservient technical role. Its function is not to question the purpose being served by schools but to improve the system's overall effectiveness in accomplishing the educational goals which have been designated by politicians and other leaders in society. One effect is that most empirical research on testing and grading tends to be superficial, and neglects most of the core issues and concerns in this area. And since it does not get at the real roots of these problems, this educational research also is basically ideological and serves the general conservative function of improving the overall efficiency with which the current educational system oppresses students.

Most of the existing research literature relating to the effects of testing and grading can be organized around the following four propositions:

1. When closely tied to meaningful occupational and lifestyle outcomes, grades can successfully serve in many circumstances as external reinforcers of learning. Grades share with such reinforcers as imprisonment, capital punishment, and money the power of diverting students' energies from their own personal concerns and problems to the expectations of outside authorities. In short, several studies show that when closely connected with primary reinforcers, grades often have the power of forcing students to learn (or at least memorize) some of the things that teachers want them to learn.

Considerable research could be cited to demonstrate that grades are often successfully employed as external reinforcers. I will limit myself to a brief discussion of one typical research study in this area. Karlins, Kaplan, and Stuart (1969), for instance, studied the performance of Princeton students in graded and pass-fail classes. At Princeton each student is allowed to take one course each year on a pass-fail basis. Instructors grade all students taking the course and it is the registrar who transforms the instructor's grade into the pass-fail designation. Students indicated on questionnaires that they learned more, worked closer to their capacity, were more motivated to learn, and more actively participated in the regularly graded courses than the ones marked pass-fail. Moreover, the researchers indicated that students received significantly better grades in their competitively graded courses than in their pass-fail courses. It seems plausible, then, that in this situation grades created sufficient outside pressures on students that they tended to withdraw energies from courses where the grading was not strictly enforced and invested themselves in the dividend-paying courses.

2. The accuracy of grades as measures of cognitive achievement is highly questionable since grades also reflect many attitudinal and non-cognitive evaluations by teachers. Rosenthal and Jacobson (1968), for instance, demonstrated that students' performance on exams could be successfully manipulated by experimenters by influencing the expectations which teachers held regarding these students. This labeling process by teachers has also been found to be influenced by such background variables as the student's family background, racial and ethnic group membership, and gender (Ohmer, 1972).

Other studies demonstrate that the grades assigned by teachers are positively related to such student personality traits as dependability, punctuality, identification with the institution, motivation by outside rewards, and the ability to follow rules. On the other hand, the grades assigned are negatively correlated with other student personality traits such as creativity, aggressiveness, and independence. As Bowles and Gintis (1976: 120-141) note in their review of this literature, many of these studies show that work organizations also reward and punish similar personality traits thus demonstrating the effectiveness of schools in socializing young persons for later authoritarian work roles.

In their widely publicized literature review, Dubin and Taveggia (1968) found no significant relationship between evaluations of student performance on final examinations and different teaching methods. Although other explanations are possible, one plausible interpretation of these results is that grades do not accurately reflect student achievement in any significant or meaningful fashion and thus it is impossible to use performance on final examinations as a criterion for measuring the effectiveness of alternative teaching methods.

There is also evidence that those students who get poor grades or drop out of college are equal in intelligence and are more creative than those who get good grades in college (Watts and Whittaker, 1968; Holland, 1960).

3. Grades in school bear little or no relationship to any relevant outside measure of adult accomplishment in society. After conducting an extensive review, Hoyt (1965) concluded that college grades show little direct correlation with adult success no matter how this success is defined. (Also see Eble, 1976: 113, for a discussion of this matter.) Thus, various studies indicate that grades in medical schools do not predict accurately one's future success as a physician, that grades in business school do not reflect accurately one's future success as a business executive, and so forth. This area of research is plagued, of course, by many measurement and statistical difficulties, but the sheer volume of evidence and the number of different indicators of success which have been employed in these studies are powerful arguments for the general accuracy of this proposition.

4. Grades suppress such student traits as originality, curiosity, questioning, open mindedness, and critical thinking, and promote such student traits as dishonesty, fear, insincerity, and blind obedience to outside authority. After reviewing a number of studies on the harmful effects of grades Robertson and Steele (1969: 61-62) summarize the major arguments against the grading system as follows:

> "1. Grades separate teacher from student. They produce coercion, manipulation, and various forms of intellectual prostitution.
> 2. Grades are inaccurate, nonuniform, and meaningless.
> 3. Grades produce a primary focus upon extrinsics and thus detract from the real purposes of education.
> 4. Grades encourage a mechanical, trivia-conscious kind of learning and prevent synoptic, penetrating thinking.
> 5. Grades stifle creativity and militate mightily toward conformity.
> 6. Grades do not motivate students effectively and they destroy the real motivation that grows out of genuine interest.
> 7. Grades are a most debilitating form of "labeling" which shuts countless students off to their own potentials as human beings.
> 8. Grades put learners in competition against each other in a very harmful way. Grading is extremely divisive and makes the spontaneous emergence of learning communities impossible.

9. Grades bring the unpleasantness of an inquisition into learning, stifling the joy and ecstasy of education.
10. Grades generate an enormous underworld of cheating.
11. Grades produce an atmosphere permeated with dishonesty, fear, and tension.
12. Grades do not, as it is commonly claimed, prepare people for life in any worthy, sensible manner.
13. Grades are not necessary for certification and employment qualification.
14. Grades are incompatible with the purposes of teaching and therefore violate the academic freedom of both teachers and students."

Grading fosters rampant cheating and institutionalized dishonesty. A survey at Columbia University by Bowers (1964) showed that in American colleges and universities over one-half of all students admitted cheating on exams. More recent studies have this percentage of student cheaters to be even higher (Ohmer, 1972). Moreover these studies demonstrate that cheating generally bears no relationship to moral standards in other areas of the students' lives. It rather seems to be generated by the pressures which schools invariably exert.

Of course, actual cheating on exams is only the tip of the iceberg with regard to the classroom dishonesty which is generated by grading. Because of these external pressures students are forced continually to wear masks and present "false selves." As Young (1975) observes: "Outside the class, the student is a lively, verbal, critical, and motivated human being. In class, the student dissimulates, frowns, masks anger and/or delight, and disassociates from the student role as fast as possible."

There is also evidence that teachers tend to bias their grades in favor of conforming students and against creative ones (Holland, 1960). Several studies have shown that creativeness generally bears no relationship to acadmmic accomplishment. Studying several major universities, for instance, Davis (1964) demonstrated that even the faculty members' own evaluation of the creativity of their students had little relation to the grades they assigned the students. Other research on the tendency of exams to penalize bright creative students and to reward less intelligent ones is summarized by Ohmer (1972). After reviewing this literature on grading and conformity, Miller (1967) concludes as follows:

> It seems reasonably clear that the grading system, at all levels, including the graduate one, tends to reward the conforming plodder and penalize the imaginative student who is likely to make a significant contribution in nearly any field. It is obvious that the discouragement and neglect that creative students tend to receive are only expressed in grades. The causes of the problem of the creative student lie deeper within our educational structure. . . . It seems self-evident that the inhospitable environment which colleges present to the creative person, and the punishment that grades visit upon him, probably do our society more harm than good.

Grading also contributes to making learning distasteful. Effort and drudgery are often glorified and students are continually exhorted to force themselves to learn. As Robertson and Steele (1969: 49) observe:

> The "good student," one is constantly reminded, is the one who puts in the long hours of "hard work," the one who "applies himself," the one who "is willing to sacrifice," who "puts his time into his studies" instead of into "frivolous" things like having fun. If you want to play music, they tell you to "practice" every day, whether you feel like it or not. The enjoyment is supposed to come later, and only after you have paid your dues in sweat, boredom, and discomfort.

There is of course no research support for this learning-as-drudgery thesis. On the contrary, considerable evidence could be offered to support the contrary thesis that creativity often is associated with voluntary, intrinsically exciting, internally motivated learning. A follow-up study of some 50 graduates of A. S. Neill's Summerhill (Bernstein, 1968), for instance, showed these adults to be living productive, happy, and varied lives. There was no evidence that these persons were in any major way handicapped by not being trained using traditional graded classroom methods.

Grading and Liberation

Below I have outlined some of my general thoughts about grading as it relates to the process of helping students to develop as self-governing, unfettered human beings.

1. As demonstrated above, reliance on grades as a relevant measure of student learning does not appear to rest on sound empirical evidence of the overall effectiveness of grades in promoting student growth. Rather, the pressures on teachers to grade students appear to originate from two general overlapping sources:

a. Schools help reproduce the class structure in the United States by providing the credentials and technical skills necessary for persons to obtain desirable jobs, the illusion of equal opportunity necessary to maintain the overall legitimacy of the hierarchical occupational structure itself, and the student attitudes and personality traits necessary for functioning in alienating and hierarchically controlled work organizations (Ewens, 1976: 308).

b. Students as products of authoritarian social relations in the schools (and elsewhere in society) develop salient needs to be controlled by outside authority. Having seldom in the past been responsible for planning their own learning and having been continually rewarded in schools for their submissiveness, dependability, and willingness to obey rules, most students quite naturally lack confidence in their abilities to govern their own education. As Friere (1970: 32-33) observes, students in this situation tend to suffer from a duality that has been established in their innermost beings. With regard to their education, they discover that without freedom they cannot exist authentically, and yet they tend to resist and fear this freedom. (I will outline below the general way this process may come about. The best general source about how this fear of freedom develops is Reich, 1970).

2. Social scientists have most often studied persons regarded as deficient in one way or another--neurotics, psychotics, persons perceived as having serious emotional or behavioral problems--rather than studying more fully developed persons at the positive end of the human development continuum. In recent years, however, there has been renewed interest in the characteristics of these more fully developed persons, and below I have summarized what I consider to be some of the most important characteristics of these persons as reflected in this literature.

a. Increasing openness to experience (versus repression): As persons develop their human capacities, they tend to move away from the pole of defensiveness and toward the pole of openness to experience. As Rogers (1961) states:

> These persons are becoming more able to listen to themselves, to experience what is going on within themselves. They are more open to their feelings of fear and discouragement and pain. They are also more open to their feelings of courage, and tenderness, and awe. They are free to live their feelings subjectively, as these exist in themselves and also free to be aware of these feelings. They are more able fully to live the experiences of their organisms rather than shutting these out of awareness.

Such living means a lessening of rigidity, of imposition of structure on experience, and of disowning feelings and experiences which are unpleasant or inconsistent with one's self image.

b. Increasing self confidence: Developing persons display a relative lack of overriding guilt, of incapacitating fears of inadequacy, and of severe anxiety. This directly contrasts with neurotic persons who are often crippled by their own feelings of guilt and inadequacy (Maslow, 1970; Perls, 1969; Rogers, 1961).

c. Increasing spontaneity and naturalness: Developing persons tend to act more often on the basis of their own perceived needs rather than on the basis of facades or masks. They put more emphasis on doing what they want or need to do in situations rather than merely trying to live up to others' expectations or trying to please others. "Self-actualizing people," states Maslow (1970):

> display simplicity and naturalness, a lack of artificiality or straining for effect. It is their impulses, thought, consciousness that are so unusually unconventional, spontaneous, and natural. In nonimportant areas, they will go through the ceremonies and rituals of convention with a good-humored shrug and with the best possible grace. But their conventionality will be cast aside when it inhibits them from doing anything that they consider very important or basic.

(Also see Perls, 1969; Rogers, 1961).

d. Healthy social relationships: "Self-actualizing persons," states Maslow (1970), "have deeper and more profound interpersonal relations than other adults. They are capable of more fusion, greater love, more perfect identification, more obliteration of the ego boundaries than other people would consider possible." Rogers (1961) adds that more fully developed persons are better able to put themselves in other's place and participate in mutual dialogue with other persons.

Based on a review of the relevant literature, I have proposed elsewhere (Ewens, 1976: 311-312) that the growth potential of social relationships can perhaps best be summarized in terms of the following four dimensions: (1) Modeling: the degree that one has the opportunity to observe a kaleidoscope of human potentials or possibilities in the relationships; (2) Leveling: the degree that one receives honest and trustworthy feedback from the other person; (3) Disclosure: the degree that one is able to express verbally on a fairly regular basis the things that one considers important, and (4) Social Support: the degree that one experiences the safety to take risks and is encouraged to trust one's own basic impulses.

e. Democratic character structure: As Reich (1945) shows, these democratic persons are more comfortable with freedom and with power equality among individuals. Whereas authoritarian persons, in Fromm's (1941) words, display sado/aschistic personalities (that is, they often behave so as to hurt others or having others hurt themselves) and function more effectively in hierarchical, dominant-submissive, social relationships, democratic persons more often prefer exchange of views, shared decisions and egalitarian social relations. (I will discuss more fully the development of Democratic character structure below.)

3. While grades are usually considered necessary for social control in the classroom they are nevertheless essentially antithetical to the development of more fully functioning persons. Instead of producing persons who are open, confident, spontaneous, democratic, and socially sophisticated, grades support authoritarian social relations in schools which produce persons who are repressed, non-confident, rigid, authoritarian, and individualistic. (Below I will essentially repeat a major part of Freud's basic argument, although Freud saw these basic problems as inhering in civilization in general rather than being generated by authoritarian social relationships.)

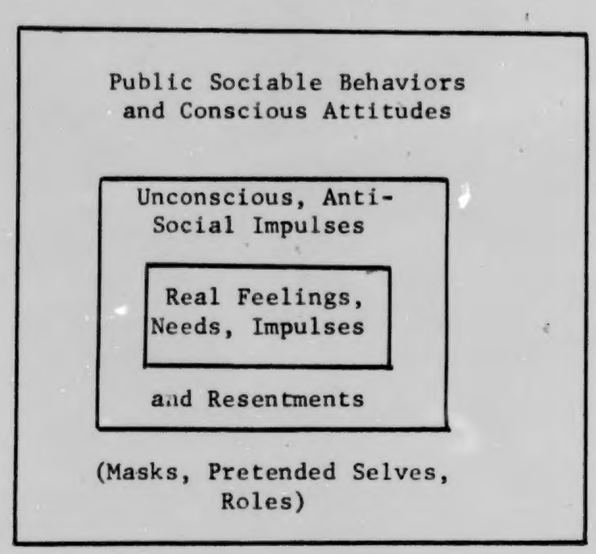
Authoritarian relationships are characterized by unilateral decision making and enforcement. In the beginning of their lives children tend to act spontaneously on the basis of immediate feelings and wants; they cry when they are hungry, smile when they are contented, and so forth. Authoritarian parents--and later authoritarian teachers--view the child's activities selectively. They encourage some activities, do not notice or are indifferent to others, and actively suppress still other activities. In the U.S., for instance, expressions of sexuality (e.g., fondling one's genitals) and natural aggressiveness are likely to be suppressed.

In the beginning, adult authorities control the child's behavior through the use of force but this is very inefficient since it requires constant surveillance. Thus, as the child develops language and memory, adult authorities are able to rely increasingly upon instilling internalized social controls. They tell the child that a good boy or a good girl (or, if they are less sexist, a good person) is one who does not hit other kids, play in the streets, play with one's genitals, crap on the floor, take off one's clothes in public, and so forth. Through the use of force children can be encouraged to anticipate their parents' expectations in this regard and some parents, of course, place considerably more emphasis than do others upon getting their children to think in terms of "shoulds" and "guilts" rather than continuing to rely on the more primitive direct control techniques. This gives rise to the child's *conscience* and the development of a conception of social *morality*. With authoritarian socialization, the conflict between one's natural impulses and society's expectations tends to move inside people's heads and they increasingly control their own behaviors by means of internally administered rewards and punishments.

Of course, one consequence of authoritarian socialization is that children often are put into untenable positions in which there are basic conflicts between their basic needs and impulses and the contradictory social expectations of authorities. In less severe instances this often leads to lying, role playing, false selves, and pretenses. In more extreme cases these contradictions may lead to *repression* of experiences in which threatening events are distorted and rendered unconscious. This role playing and repression of course requires considerable energy which is not productive with regard to developing a person's capacities and potentials. In particular, repressed experiences have to be constantly shut out of consciousness and the resulting symptoms usually cause the person secondary types of problems (e.g., fears, phobias, compulsions, anxieties). These repressed experiences also cause a basic split between a person's outward behaviors and consciously held beliefs and the person's real feelings, needs, and impulses. As illustrated below, persons in authoritarian environments tend to develop an intermediate layer of personality which is composed of unconscious, anti-social impulses and resentments, and this intervening personality layer forges a basic separation between persons' outward public behaviors and attitudes, on the one hand, and their own internal core impulses and feelings, on the other hand.

The result of authoritarian environments, thus, is internally controlled, repressed persons who must constantly monitor and suppress their own basic impulses. In contrast to the more fully developed persons described above, people who are products of authoritarian environments are characterized by much guilt, disowned parts of self, and resentments toward authority. They learn to exist in a basic condition in which their public selves become more or less permanently estranged from their own core feelings and needs.

They also come to function in a society where this basic split condition is also the natural state of other persons with whom they must deal.


Public Sociable Behaviors and Conscious Attitudes
Unconscious, Anti-Social Impulses
Real Feelings, Needs, Impulses
and Resentments
(Masks, Pretended Selves, Roles)

In authoritarian environments, persons also tend to become separated from authorities. The developing child is put in the uncomfortable position of depending on the same persons (e.g., parents, teachers) who also directly or indirectly (through conscience) suppress many of their own basic impulses and activities. This naturally leads to a basic split between the individual and society (e.g., id-superego, individual-civilization, student-teacher) which is basically unreconcilable. Within the individual person this split is reflected as a deep ambivalence toward one's parents, teachers, and other authorities (and authority in general), on the one hand, and a basic fear of freedom (as described above), on the other hand. This ambivalence toward powerful others is of course extremely uncomfortable and many persons are motivated either to repress their dependency feelings (Freud's explanation of radicalism--infantile rebellion against the authority of the father) or to repress their hostile feelings (and over identify with authority--Reich's Authoritarian Personality).

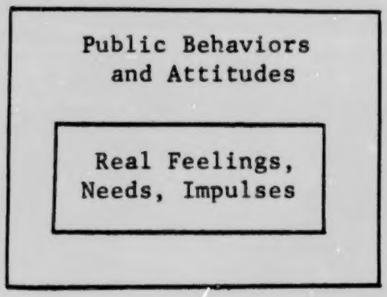
The characteristics of more fully developed persons that I have described are not nurtured in authoritarian environments. Such environments do not lead to deep or healthy social relationships. Because power is relied on as the basis for joint activities there tends to be little transparency in these relationships. One cannot afford to disclose one's true feelings, give honest responses, or take risks when mistakes are likely to be punished. Thus, in terms of the growth potential of social relationships there tends to be little disclosure, leveling, social support, or modeling. And persons in these dominance-submission relationships tend not to develop the deep trust, intimacy, warmth, and empathy necessary for more full human development.

Since persons do not learn to take basic responsibility for their own lives and become self-regulating they cannot be relied on to govern their own behavior. Beneath the fragile facade of sociability and culture--as every conservative knows only too well--is a destructive beast poised and ready to escape (the intermediate layer of personality in the illustration above). Just as domesticated rabbits, after generations of dependence on others for nurturance, can no longer take care of themselves in the wild, so these domesticated persons, who have been raised to be governed by outside authority, can no longer be counted on to avoid socially destructive and self destructive behaviors.

4. It is a natural tendency for persons raised in the context of authoritarian relationships to conceptualize the basic alternative as permissiveness. But permissive relationships are only a negation of authoritarian relations and thus tend to be power relationships in which only the respective dominant and submissive parties have changed places. I propose that transcendence of this authoritarian-permissiveness dimension should be through democratic relationships in which decisions are reached on the basis of mutual acceptance and complied with because of self interest. And since there is little or no suppression by parents and other authorities, individuals are seldom placed in untenable situations of fundamental conflict between their own basic impulses and contradictory outside social demands. Thus, in general there is no need for pretense or the presentation of false selves. Persons have veto power over basic decisions which affect their lives and the security of knowing that their own judgments cannot routinely be overruled.

By the same token, there is little need for persons in democratic relationships to repress unpleasant experiences. Since outside authorities do not presume to have the basic power to control and redirect one's natural sexual and aggressive activities, for instance, individuals are not placed in the precarious position of having to monitor and control rigidly their own spontaneous and natural impulses. On the contrary, since agreements are always negotiated, people's power and ability to control the forces which shape their lives are directly related to their ability to know and understand their own real feelings, needs, and impulses in given situations.

Suppose that students are in a democratic classroom, for instance, in which they must come to a mutual agreement about readings, course procedures, evaluation techniques, and so forth. The ability of given students to use their time wisely in this class will in large measure depend on their knowing why they took the class in the first place and what basic questions and problems in their lives can profitably be dealt with in this course. Thus, survival in democratic relationships requires persons who are socialized to understand their own core feelings and needs. Rather than developing an intermediate layer of unconscious, anti-social impulses and resentments which intervene between public behaviors and core feelings, democratic relationships should promote personality development in which one's public and private selves are closely related to one another. Democratic environments should thus promote the development of a simpler two layer personality structure as presented below:


Public Behaviors and Attitudes
Real Feelings, Needs, Impulses

Since persons do not develop this destructive intermediate layer of personality they tend to develop the basic characteristics of more fully developed persons as described above: openness to experience, self-confidence, spontaneity, healthy social relationships, and democratic character structures. In short, democratic relationships rather than the domestication of individuals tend to produce persons who are self-regulating and who have the experience and the capability to govern their own lives.

There is also little basis for resentment against authority since the individual is at the same time less dependent and less vulnerable to powerful others and also seldom suppressed by them. Social relationships thus tend to develop which are characterized by high degrees of leveling, disclosure, social support, and modeling. Democratic relationships tend to generate situations in which persons are mutually interested in one another's development and in which a threat to one is often perceived as a threat to all. Thus if my own personal well being is vitally linked to the response and support which I am receiving from another person, then the things which hurt that other person will tend to be perceived as hurting me as well.

External social controls are also less necessary in situations involving unfettered, self-governing persons. Since there is no basic split between their public and private selves and no fundamental estrangement from external authority, these individuals can usually be relied on to act in an autonomous and socially responsible manner. There will of course be struggle and conflict but the resolution of these difficulties will tend to lead to personal growth and to group solidarity rather than to their destructive opposites. Interdependence and community are basic facts of human life and democratic relationships allow these fundamental realities to emerge and dominate the consciousness and behavior of individual human beings.

5. Teachers who are committed to help students develop as free, unfettered, and self-governing persons are thus caught in a number of fundamental dilemmas regarding grading. On the one hand, there are strong external pressures which require grading. The authoritarian system requires student grades to efficiently reproduce itself; and students who are products of that authoritarian system have also developed both strong needs for outside controls and a basic ambivalence toward taking responsibility for their own lives. The situation is of course further complicated by the fact that we as teachers also are products of the same authoritarian system and also tend to function most effectively in the context of hierarchical relationships.

Moreover, with regard to schools as social institutions, teachers are caught in the basic reform dilemma which is best described in another context by Gorz (1964). To help students achieve autonomy and freedom requires democratic relationships within schools which would only be possible on the basis of a total restructuring not only of formal education but of other major institutions in American society as well. But to deal with the problems related to grading in piecemeal and reformist ways as most researchers in this area propose (for instance, Eble, 1976; and Dressel et al., 1961) will merely further streamline and make more effective the system's present capabilities for oppressing students.

Teachers are thus caught in a fundamental bind. On the one hand, they are not themselves in the position of bringing about a total revolutionary restructuring of society. But merely waiting for this event to take place by itself is probably also a self-defeating tactic. Having not personally engaged in the struggle to change society, persons will probably have neither the capabilities nor the consciousness to take advantage of this event should it occur. But engaging in clearly reformist activities, as is argued above, is also likely to have the effect of strengthening the present oppressive system and of making it more difficult for students to gain ultimate control over their own lives.

The way out of this reform dilemma, as Gorz (1964) suggests, is to distinguish between reformist and non-reformist reforms. Reformist reforms, as described above, have the effect of strengthening the prevailing logic of the present oppressive system. Non-reformist reforms, on the other hand, challenge the prevailing authoritarian social relations in fundamental ways by giving persons more direct control over their own lives. Thus, with regard to grading, teachers must individually and collectively struggle for reforms that will allow students to have more control over their own education and the institutional context and process by which this education takes place.

Below I will describe my initial thoughts about some of the things which I would consider to be non-reformist reforms as these relate to the grading of students:

1. As professional educators we must start by insisting in the U.S. that formal schooling and job certification be separated. The goal of education must be to create self-regulating, critical thinking, fully-developed human beings and not to train students for jobs. When students must be certified by their teachers before they are deemed qualified for a job then in the present society authoritarian social relations in schools become a

necessity. As Robertson and Steele (1969: 61) conclude: "Teachers can't be inquisitors, and when it's up to teachers to say who is qualified for certain jobs, they are inquisitors. There is no escaping it."

Given the continuing process in advanced capitalist societies of increasingly cheapening and degrading work (see Braverman, 1974), perhaps 80 percent of the work force is already over educated for the actual work which they perform (Berg, 1971). Thus also at a practical level justification of formal education in this society on the basis of human development is easier to make than on the basis of training for necessary job skills.

2. Following the basic logic of scientific management and capitalist industrialization, schools also are becoming increasingly managed institutions with growing control over educational processes being appropriated by central administrators. Thus if democratic social relations are to dominate in schools this trend must be reversed. Colleaguality and democratic decision-making must be restored and expanded in schools.

As emphasized above, schools are never neutral in this regard. Either they are governed by hierarchical authoritarian social relations with basic control centralized at the top, in which case they become factories for domestication and the production of students with authoritarian character structures, or through democratic processes they become nurseries for the growth of self-governing and free human beings (Friere, 1970).

3. As Robertson and Steele emphasize above, grades are incompatible with the ultimate purposes of the teaching-learning process and thus violate the academic freedom of both teachers and students. Instruction must have as one of its goals making students aware of the contradictions in their lives which are produced by grades. Thus, we must encourage student consciousness of the many ways in which authoritarian relations in the classroom demean and domesticate them and ultimately destroy their abilities to become unfettered, self-regulating human beings. Our ultimate responsibility as teachers of course is to help students be able to control their own lives. This means enlarging areas in their lives over which they have direct responsibility and helping them gain experience in making mutual, democratic decisions.

Learning Sociology and Assessing Critical Thinking*

Paul J. Baker
Illinois State University

Introduction

Sociological inquiry rests on the noble claims of human reason (Bendix, 1970). Sociologists who make their living in the academy often assert this commitment to reason in the company of undergraduate students. Sharing the educational goals of their learned colleagues, sociologists join the chorus of academicians who believe that students "should be able to think clearly, to read critically, and to write and speak effectively" (University of Illinois, 1979: 6). The rhetoric of curriculum committee reports and college catalogs is also supported by a recent national survey of sociology departments. According to Bradshaw and McPherron, chairpersons in universities, four-year colleges, and community colleges agree that the most important instructional goal in the sociology curriculum is the development of critical thinking (1978: 17).

Emphasis on critical thinking in sociology raises some interesting questions. If sociologists have placed critical thinking at the top of their curricular goals, why has the topic been ignored in their professional literature on the assessment of learning outcomes? If critical thinking is so important, why have so few sociologists designed testing methods which can demonstrate the achievement of such educational aims? These questions are rhetorical, but provoke an opportunity to explore uncharted territory in the field of sociological instruction. We have little systematic knowledge about the testing of learning outcomes in sociology courses; even less information is available on the topic of learning sociology and assessing critical thinking.

This chapter is divided into two sections. First, knowledge available on the topic of critical thinking as related to sociology will be explored. While sociologists have spent little time on the topic, they can benefit from several theoretical and empirical investigations conducted by educators and philosophers. After reviewing knowledge available, I will discuss crucial areas in which knowledge is needed to improve the sociology teacher's capacity to assess critical thinking.

Knowledge Available: Styles of Teaching and Testing Techniques

Testing critical thinking is not an isolated instructional task. Rather, it is one component of a larger set of teacher-student interactions. This dyadic relationship is often examined by educators according to types of teaching styles and types of testing procedures. During the past half century, many educators have described a common pattern in which certain types of instruction are associated frequently with particular methods of evaluating learning outcomes (Beyer, 1971; Chickering, 1976; Ausubel, 1961). In its simplest form, the educators' model is bi-polar. Expository teaching is related to objective exams in which students are expected to replicate the exposition of textbooks and lectures. Critics of expository teaching charge that it involves little more than listening, reading, copying, and memorizing discrete recall items conveniently designated "T," and "F," or "A," "B," "C," and "D." At the other extreme, educators describe the inquiry approach; in this setting, students engage in problem-solving tasks which challenge their capacity to extend the realm of school knowledge to new areas outside the classroom. William Gartner has described these teaching/learning strategies as tell-recite contrasted with excite-discover (1970).

Recognizing the conceptual crudeness of this bi-polar typology, educators often stress the fact that teaching techniques can be placed along a continuum. The degree of elaboration varies, but educators often mention one or two intermediate types (Bigge, 1976: 327-28; Chickering, 1976: 90-91). Most typologies hold clear hierarchical learning concepts that follow Bloom's taxonomy of educational objectives or Gagné's eight conditions of learning. Students climb a progressive ladder of sophistication from the primitive memory skills of recall to the sophisticated writing tasks of analysis, evaluation, and problem solving.[1]

Academic folklore suggests that objective exams are concerned only with the assessment of fact retention; conversely, essay exams are always aligned with the evaluation of critical thinking. Such impressions from common sense and bi-polar typologies must be examined with the keen eye of qualifying exceptions. Some educators construct critical thinking tests in an objective format, and many teachers administer essay exams which require little more than direct replication of information found in textbooks and lecture notes. Teachers who devote their time to reading essay exams are not necessarily evaluating the critical thinking of their students. In short, critical thinking is not associated with any single pattern of teaching and testing.

The educational goals of critical thinking can never be divorced from the broader context of student-teacher expectations; but, it is difficult to determine the occurrence of critical thinking by merely examining various types of teaching and testing techniques. The problems of placing critical thinking into a broader learning context must await clarification of the term itself. It is time to examine the nature and meaning of critical thinking.

Knowledge Available: Diverse Definitions of Critical Thinking

Like many important terms (e.g., scientific method, sociology), critical thinking defies precise definition. The vagueness of the term, however, does not nullify valuable work by educators and philosophers who have identified a set of intellectual skills essential to the learning process both within and outside formal school settings. I shall begin the task of definition by discussing some mental activities which are not included under the general rubric of critical thinking. While scholars have always found imaginative and creative thinking worthwhile, these thought processes often lack sufficient logical rigor to be included as part of critical thinking. Many intellectual abilities identified by Bloom at levels 1, 2, and 3--knowledge, comprehension, translation, interpretation, extrapolation, and application--would also be excluded from any list of critical thinking skills. These skills are important, but considered to be too elementary. There are many complex thinking processes which require more effort than rote memory yet are not included in various lists of critical thinking abilities.[2] Finally, the intellectual tasks associated with symbolic logic are rarely included in the province of critical thinking.[3] Some contemporary logician-educators link critical thinking with the informal logic movement of recent years (Ennis, 1979).

No simple schemes delineate all topics which have surfaced as aspects of critical thinking. Nevertheless, I shall clarify the subject by suggesting that educators have adopted one of two approaches. First, following the inspiration of John Dewey, educators have defined critical thinking as a problem-solving process. This approach is sometimes identified by science and social science teachers as the discovery method. Second, other educators have acknowledged Dewey's contribution, but have made an extension to include additional topics of deductive logic and argument analysis. Some of the thinking skills developed by educators in these two loosely-defined learning orientations follow.

Dewey made a sharp distinction between the process of reasoning and the study of formal logic. Logical forms are not used in actual thinking, but merely set forth the results of thinking.

> [Logical] forms apply not to teaching conclusions, not to arriving at beliefs and knowledge, but to the most effective way in which to set forth what has already been concluded, so as to convince others . . . of the soundness of the result (1933: 74).

According to Dewey, "Actual thinking has its own logic; it is orderly, reasonable, reflective" (1933: 75). He believes that every thinking experience begins with "a perplexed, troubled, or confused situation"; it will lead to a "cleared-up, unified, resolved situation at the close." Coining the term "reflective thinking," Dewey summarized five phases of this cognitive process (1933: 107):

(1) *suggestions*, in which the mind leaps forward to a possible solution;
(2) an intellectualization of the difficulty or perplexity that has been *felt* (directly experienced) into a *problem* to be solved, a question for which the answer must be sought;
(3) the use of one suggestion after another as a leading idea, or *hypothesis*, to initiate and guide observation and other operations in collection of factual material;
(4) the mental elaboration of the idea or supposition (*reasoning*, in the sense in which reasoning is a part, not the whole, of inference); and,
(5) *testing* the hypothesis *by* overt or imaginative *action*.

These five phases are not necessarily adopted in a fixed order. Different problems will involve different patterns with some phases overlapping, reoccurring, expanding in importance, or diminishing in significance. Dewey certainly did not intend to create a mechanical method of thinking which could be uniformily divided into five phases.

The irony of Dewey's philosophy is the ease with which his pragmatic ideas became textbook dogma in the potboiler literature produced for college students planning careers in American public schools. Thousands of students memorized the five stages of reflective thought as gospel truth. But Dewey's model acted also as a stimulus and guide for college teachers committed to the educational aims of general education (Dressel and Mayhew, 1954).

Many philosophers and educators believe Dewey too extreme in stressing the contrast between formal logic and reflexive thinking. But, they still accept his concept of critical thinking as a cognitive process. Critical thinkers approach a given problem or statement with a coherent set of steps or stages which can be employed flexibly as a unified system of logical operations. Elements of this cognitive process are mobilized to attack a statement or problem as a total thinking experience. The problem or statement is seen as a whole unit of intellectual inquiry. Critical thinking involves a relatively short list of mental activities essential to understand the thoughts of others and to construct a thoughtful position of one's own (Dressel and Mayhew, 1954; Angel, 1964; Scriven, 1976).

A second approach to defining the mental skills of critical thinking is more complex and elaborate. Philosophers of this orientation are not interested in quarreling about various schemes of reasoning or reflexive thinking. Rather, they extend the definition of critical thinking by adding numerous topics of logical analysis. Max Black's widely-adopted text, Critical Thinking (1952), presents an array of topics including deductive logic, language usage, induction, and the scientific method. Black's work is instructive, but the taxonomy on critical thinking constructed by Robert Ennis is the most comprehensive and systematic statement extant. Defining critical thinking to be "correct assessing of statements," Ennis articulates twelve aspects of this elusive topic (1962: 84):

1. Grasping the meaning of a statement
2. Judging whether there is ambiguity in a line of reasoning
3. Judging whether certain statements contradict each other
4. Judging whether a conclusion follows necessarily
5. Judging whether a statement is specific enough
6. Judging whether a statement is actually the application of a certain principle.
7. Judging whether an observation statement is reliable
8. Judging whether an inductive conclusion is warranted
9. Judging whether the problem has been identified
10. Judging whether something is an assumption
11. Judging whether a definition is adequate
12. Judging whether a statement made by an alleged authority is acceptable.

While this list is comprehensive, it does not exhaust all aspects of important and sophisticated thinking. For example, Ennis excludes by definition any discussion of creative thinking; and he excludes temporarily the judging of value statements because the topic is too complicated and controversial. Ennis recognizes that topics on the above list are not mutually exclusive. And finally, he stresses the importance of pragmatic judgment in the assessment of any statement. The pragmatic dimension "requires the admission that complete criteria can not be established for critical thinking. An element of intelligent judgment is usually required in addition to applying criteria and knowing the meaning" (1962: 85).

During the past seventeen years, Ennis has refined, extended, and tested empirically many items in this original taxonomy (see especially, 1964, 1969, 1973, 1975). He is preparing currently a revised statement on critical thinking more extensive than the original. Re-naming his area of inquiry rational thinking, Ennis elaborates on his original twelve topics. He outlines briefly characteristics of rational thinkers as persons who:

A. Are *proficient* at:
   1. *observing*;
   2. *inferring* an explanation, the point, the meaning, etc.;
   3. *generalizing*;
   4. *conceiving and stating* assumptions, alternatives, plans, predictions, definitions, etc.;
   5. *offering* a well-organized and well-formulated *line of reasoning*;
   6. *evaluating* authoritative-sounding statements, deductive reasoning, inductive reasoning, explanations, value statements, and definitions; and
   7. *detecting standard problems* and *realizing appropriate action*; and

B. Have the *tendency* to:
   1. *exercise the proficiency* they possess;
   2. take into account the *total situation*;
   3. be *well-informed*;
   4. demand *as much precision* as the subject matter permits;
   5. deal with the parts of a *complex situation in an orderly fashion*;
   6. consider seriously *other points of view* than one's own;
   7. *withhold* judgment when the evidence and/or reasons are insufficient;
   8. *take a position* (and *change the position*) when the evidence and reasons are sufficient to warrant so doing; and
   9. accept the necessity of *exercising informed judgment*; and

C. *Exercise good judgment* (1979: 5-6).

In 1962, Ennis defined critical thinking as a distinctly *reactive endeavor* in which the critical thinker responded with rigorous scrutiny to the thoughts of others. Ennis has shifted his position in recent years, presently defining critical thinkers as more positive and assertive persons. They are expected not only to detect a variety of problems, but to take appropriate action. Critical thinkers actively seek to become well informed, and do not shy away from value questions.

Knowledge Available: Some Perennial Issues and New Terminology in the Field of Testing

Sociology teachers can benefit from closer examination of various concepts of critical thinking. They also can gain new insight on their teaching craft by becoming familiar with current literature about evaluating learning outcomes. During the past two decades, conventional ideas of testing have been challenged among educational researchers. The controversy has centered on the purposes for which tests are intended. One such debate distinguishes between criterion-referenced tests and norm-referenced tests. Other scholars identify the functions of testing by using the contrasting terms of normative and summative evaluation. I will review briefly these basic taxonomies.

During the past half century, testing technicians have developed discriminating tests which rank people according to group norms (the group may be a classroom, school, state, or the nation). Tests such as the SAT, ACT, and GRE are designed to make comparative judgments among people. They serve as sorting devices to determine placement in various learning settings (e.g., "slow" classrooms, elite universities). While critics often have denounced such tests, an alternative measurement theory did not surface until 1963. In a landmark article, Robert Glaser introduced a new dichotomy. "What I shall call criterion-referenced measures depends upon an absolute standard of quality, while what I term norm-referenced measures depends upon a relative standard" (1963: 519). Norm-referenced tests determine a student's relative standing (e.g., 90th percentile); criterion-referenced tests determine a student's level of proficiency along a continuum of achievement (e.g., highly proficient). Criterion-referenced tests "provide information as to the degree of competence attained by a particular student which is independent of reference to the performance of others" (Glaser, 1963: 519). The issue for Glaser is not to weigh the relative worth of one's scores, but to demonstrate the mastery of a specified content area.

Glaser's concern is hardly new. Criterion-referenced measures--e.g., tests for a driver's license, a lifeguard certificate, Boy Scout Badges--have been institutionalized for many years. But, his insistence on using criterion-referenced measures in classrooms seemed revolutionary to many technicians in the testing industry.

Educational researchers such as Glaser and Popham are pursuaded that the full range of academic subjects in schools is best evaluated by criterion-referenced tests. Others disagree, causing the topic of testing to be highly controversial (Ebel, 1978; Popham, 1978). Perusal of educational literature suggests a clear trend favoring future development of sophisticated criterion-referenced tests (Hambleton et al., 1978). In many respects, the debate is not over types of tests but rather the functions tests serve. Tomko and Ennis have clarified this issue with an important recommendation (1978: 11):

> Since it appears that the same test can be used both as a criterion-referenced test and a norm-referenced test, we propose to change the labels slightly and talk of criterion-referenced testing and norm-referenced testing. This labeling explicitly recognizes the dependence of the distinction upon purpose and interpretation in the given situation. This relabeling relieves us of the burden of classifying every test as one or the other type, a task we found in practice to be impossible.

The second taxonomy focuses on appropriate occasions of testing during the instructional cycle. Coining the terms "formative" and "summative" evaluation, Michael Scriven (1967) clarified different uses tests can serve during different time periods. Evaluation with feedback opportunities during the developmental process is strikingly different than evaluation as the final product of this process. His insight was amplified by Bloom, Hastings, and Madaus in their classic work, Handbook on Formative and Summative Evaluation of Student Evaluation (1971). These authors added a third type of test--diagnostic--to Scriven's original typology.

Diagnostic testing assesses the student's level of knowledge and learning capacities either prior to or at the beginning of instruction. Such tests serve several purposes. They may record previous achievement, determine presence or absence of prerequisite skills, or classify the student according to various characteristics considered relevant for alternative modes of instruction. In some cases, diagnostic instruments serve as pre-test data, establishing a baseline to measure the intervention effects of an instructional experiment. Finally, diagnostic tests can be shared with students as an early feedback system which helps them assess strengths and weaknesses in their own learning. Diagnostic tests serve both formative and summative evaluation purposes.

The main purpose of formative evaluation is to help students and teachers assess the learning outcomes while the class is still in the midst of its work. "The purpose is not to grade or certify the learner; it is to help both the learner and the teacher focus upon the particular learning necessary for movement toward mastery" (Bloom et al., 1971: 61). Summative evaluation, on the other hand, provides the final overall assessment of the students' learning achievements. It is the sum total of the student's recorded accomplishment for a given unit of instruction.

Glaser's dichotomy between criterion-referenced tests and norm-referenced tests is not logically inconsistent with Scriven's distinction between formative and summative evaluation. Formative evaluation is always criterion-referenced measurement; summative evaluation may be either criterion- or normative-referenced. This brief excursion into the jargon of educational testing will be useful during later discussion of various tests of critical thinking. The functions of testing can best be understood through a concrete assessment of empirical instances in which teachers and researchers have articulated explicitly the educational goals of critical thinking. It is time to move beyond conceptual discussion into an examination of empirical studies about critical thinking and the teaching/learning process.

Knowledge Available: Empirical Studies of Critical Thinking

Empirical scholarship on the topic of critical thinking first emerged in professional literature during the 1930s. Some of these early studies are still worthy of careful scrutiny. To this day the pioneering investigations of Edward Glaser (1941) and Ralph Tyler (1942) offer something of value to sociologists. During this same time period, sociologists investigated questions of critical thinking by studying educational methods capable of counteracting propaganda. Ironically, their educational research failed to generate empirical studies of enduring value. As researchers of human behavior, the educators seemed more sophisticated.

Relevant Studies of the High School Curriculum. As early as 1932, Ralph Tyler began to measure systematically the learning outcomes of college instruction. Guided by his interest in "higher mental processes," he studied correlations between the students' ability to recall information and their abilities to apply principles to new situations and to draw inferences from data not presented earlier. His statistical analysis led to convincing conclusions: "Memorization of facts frequently fails to result in the development of higher mental processes. If the higher mental processes of application of principles and inference are really to be cultivated, learning conditions appropriate for their cultivation are necessary" (1936: 17).

During the Depression, more young people remained in school, forcing administrators to cope with a more heterogeneous student body. In response to the new problems facing the high school curriculum, The Progressive Education Association launched a major eight-year experimental project with thirty schools throughout the United States. Tyler directed the evaluation staff for the Eight Year Project. His early empirical work with college students made him ideally suited to develop new learning assessment instruments for an expanding educational mission. Among Tyler's many achievements in this project was the creation of several critical thinking tests: Test 1.41--Social Problems; Test 2.51--Interpretation of Data; Test 5.11--Nature of Proof (Smith & Tyler, 1941). Since the Interpretation of Data test was used later by sociologists studying the effects of sociological instruction (Cook & Koeninger, 1939; Cook, 1950), it will be described briefly.

An important intellectual skill promoted in formal schooling is the ability to interpret various types of data such as news items, advertisements, scientific reports, public opinion polls. Tyler and his associates reasoned that students can appraise an information item with one of four possible responses: (1) an accurate interpretation, (2) an overgeneralization, (3) an undergeneralization, and (4) crude errors. These responses were used as a guide to study data presented in various forms such as tables, prose, charts, and graphs. A jury of experts constructed fifteen interpretive statements for each information item. Any given statement could be scored as true, probably true, false, probably false, or indeterminate because of insufficient data. By matching student responses with expert opinion, Tyler constructed a complex tally sheet with two global scores (general accuracy; crude errors) and five subscores (agreement with experts on true/false; probably true/probably false; insufficient data; complex measures of overgeneralization; and, undergeneralization). The Interpretation of Data is imaginative. But the scores cannot be reduced to simple numerical indices of critical thinking abilities. Furthermore, Tyler warns that it is "inadvisable to interpret scores on this test in terms of national norms" (1941: 60).

The second classical empirical study of critical thinking is Edward Glaser's doctoral dissertation, An Experiment in the Development of Critical Thinking (1941). In collaboration with Professor Goodwin Watson, Glaser developed a battery of critical thinking tests which were administered as pre- and post-test instruments in four experimental classrooms and in four control classrooms. Some of the curriculum materials used in the experimental classes were developed by the Institute for Propaganda Analysis. Glaser's measuring instrument was an extensive revision of the Watson tests of fair-mindedness published in 1925. The Watson-Glaser Tests of Critical Thinking contained several measuring devices: Test A1-A2--A Survey of Opinions; Test B-AR--General Logical Reasoning; Test 3--Applied Logical Reasoning Test; Test C--Inference Test; Test D--Generalization Test; Test E --Discrimination of Arguments; Test F--Evaluation of Arguments. Glaser used other tests such as the Nelson-Denny Reading Test, and a variety of observational techniques. In Denzin's terms, the investigation is an example of triangulation (1970: 26). With elaborate data and cautious optimism, Glaser reported that a self consciously constructed curriculum could facilitate the learning of critical thinking skills. The experiment was successful.

During the 1930s, sociologists and educators were alarmed by the persuasive power of propaganda. In collaboration with the Institute for Propaganda Analysis, Alfred McClung Lee and Elizabeth Briant Lee made extensive studies of propaganda techniques; they were convinced that new kinds of educational literature could help students "become more alert to the tricks of the propagandist's trade" (1939: 132). They urged the reader of The Fine Art of Propaganda to "Beware of your own prejudices. Suspend your judgments until more sides of the issue are presented. Analyze them" (1939: 134). Public schools often used Institute literature, and teachers offered testimony regarding its effectiveness. Despite these testimonies, numerous empirical studies were conducted with no conclusive results (Doob, 1950: 450). Many of the studies concentrated on attitude changes, and none of them defined carefully the nature of critical thinking. The most rigorous study using a quasi-experimental design failed to demonstrate the effectiveness of anti-propaganda instruction (Osborn, 1939: 1-17). Numerous sociologists and social psychologists wrote extensively about fascism, authoritarian personalities, and propaganda, yet a review of this literature fails to yield any new insights about critical thinking, how it might be taught and learned, or how it might be measured.

Following Sputnik and the publication of Bruner's *Process of Education* (1960), the intellectual ferment in American public schools reached new levels of intensity. During the 1960s and early 1970s, educators and social scientists led a curricular reform movement known as "The New Social Studies" (Haas, 1977). As a result, new teaching/learning materials were produced from over one-hundred special curricular projects. Many of these projects paid special attention to issues of critical thinking. Given such extensive interest in more sophisticated learning opportunities, it seems logical to ask: did this multi-million dollar educational movement generate any new knowledge about learning and about appraising critical thinking? After extensive review of this literature, I have concluded that little new knowledge was gained. But *little* is better than *none*, and two of these projects deserve brief mention.

Under the direction of Robert Angell and his staff, many prominent sociologists helped to develop an extensive high school sociology curriculum entitled "Social Resources for the Social Studies." They produced twenty-three "Episodes in Social Inquiry" which dealt with such diverse topics as "Roles of Modern Women," "Delinquency," "Simulating Social Conflict," and "Social Mobility in the United States." Everett K. Wilson articulated the project's basic rationale for inquiry (1970: 120):

> I've been suggesting that a stress on inductive learning (1) promotes participation, (2) constrains us to consider the relevance of problems posed, and (3) promotes a commitment to the task not so readily engendered by conventional teaching patterns. Participation, relevance, and task commitment, in turn speed learning and lengthen retention span.

This problem-solving approach, with a strong emphasis upon inductive reasoning, was brought to its logical conclusion on occasions of evaluating student performance. To quote from the instructor's manual: "SRSS does not advocate" essay tests or objective tests. These evaluation devices "too often suggest that knowledge is at a standstill" (Sjoberg, 1973: 52). The preferred methods of evaluation include problem-solving exercises, miniature empirical studies, and group discussion.

During the developmental phase of the project, extensive testing as formative evaluation was conducted by F. Lincoln Grahlfs. Despite considerable pressure from the publisher, Angell refused to allow these formative instruments to go on the market as summative devices. SRSS Episodes contain neither glossaries of sociology terms nor test items which measure the students' capacity to memorize such terms. The SRSS project was committed to producing high quality curricular materials which would not compromise the principles of inquiry. Unfortunately, no research findings on the inquiry model were ever published from the project. It is impossible to know if the learning outcomes hypothetically anticipated could ever be empirically demonstrated.

The SRSS materials cannot be dismissed as elementary high school literature. They demand far more rigor than the typical basic college course and are the most systematic and comprehensive sociology curriculum ever constructed on behalf of the inductive approach. In my judgment the major drawback to the SRSS material, as college literature, is its labor-intensive character. Most teachers of the basic curriculum in higher education do not want to work that hard with ambiguous inquiry techniques.

The intellectual achievement of the Harvard Social Studies Project is impressive also. Promoting a new concept of citizenship education, Oliver, Shaver, and Newmann developed a sophisticated critical thinking model. They assume that the United States is a diverse society with numerous value conflicts (e.g., individual autonomy vs. community welfare, majority rule vs. minority rights). A series of public issues such as school desegregation is assessed analytically in terms of various value conflicts. Labeling this teaching/learning method the jurisprudential approach, the authors attempted to help students rationally evaluate public controversies. Their framework is not a superficial discussion of current events in which "one's opinion is as good as another's." Emphasizing inquiry and dialogue, the Harvard materials demanded highly disciplined skills of critical thinking.

The Harvard Project had both a developmental and research orientation. For purposes of this chapter, the most important aspect of the Harvard Project is an empirical evaluation of critical thinking tests. Oliver and Shaver found inadequate all available tests of critical thinking. Their chief criticism is the validity of existing instruments to measure complex learning goals of citizenship education. The structured nature of multiple-choice tests often appears unrealistic as they tend to measure only fragments of critical thinking. Current critical thinking tests give little consideration to "the components of the total decision-making process in public controversy" (1966: 183). "In real life we don't sit at a desk checking or writing responses, but are often engaged, as part of the political process, in dialogue with others" (1970: 487). However, current paper-and-pencil tests measure only specified logical operations, without placing these mental skills in a societal context in which issues are debated. "In short, the tests tend to be unlike the context in which students are likely to apply the conceptual framework required to clarify and judge pub[illegible] issues" (1966: 184).

The validity problem must be explored in terms of evidence on the usefulness of tests in predicting critical thinking skills outside the testing situation. Is it possible that a student would score highly on a paper-and-pencil test, "yet be totally inept at performing the same operations in a 'real' setting?" (1966: 185). Oliver and Shaver attempted to answer this question by first constructing a reliable evaluation instrument to measure critical thinking competence in discussion settings. Just over one-hundred students participated in the "crucial experiment"; scores generated in the discussion setting were compared to scores obtained from pencil-and-paper tests. The correlations were extremely low, indicating that "there may be little or no relationship between competence to defend one's point of view in public and competence to do well on available 'critical thinking' tests" (1966: 225). While this study of one-hundred students hardly constitutes a major educational experiment, it does point in the direction of much needed research. In a small way, the study substantiates Oliver and Shaver's reservations about standardized critical thinking tests which claim to measure significant learning outcomes in citizenship education.

Space does not permit elaborate discussion of the Harvard Project's analytical model of instruction or the experimental design used to test its educational effectiveness. Suffice to say that the researchers used pre-post test data, and experimental and control groups. Research results from this quasi-experimental format supported their claim of significant gains for students using the Harvard Project materials. In contrast to Glaser's findings (1941), students in the experimental classes did not show significant improvement on the Watson-Glaser test. It is difficult to interpret research results in educational experiments on critical thinking. After reviewing several studies, I am in agreement with Oliver and Shaver; generalized tests of critical thinking intended to measure transfer of learning skills may show little or no significant differences between experimental and control groups. But when students are coached to develop specific types of reasoning exercises contained on the tests, they respond to the coaching by yielding significant experimental results. Essentially, the teaching/learning experiment is little more than coaching for the exam. Research literature on critical thinking is dominated primarily by studies of high school students.

Relevant Studies at the College Level. I have discussed some of the more important studies in this area. Attention will now turn to the work of sociologists at the college level. Published literature on this topic is exceedingly sparse. During the past forty years, only two sociologists have attempted to assess learning outcomes in terms of critical thinking skills.

Using Tyler's "Interpretation of Facts Test" as a measure of critical thinking, Cook and Koeninger collected pre- and post-test data on sixteen elementary sociology classes at Ohio State University (1939). The average net learning gain for all seven sub-scores was negligible. Disappointed in these test results, Cook conducted a second study with educational-sociology classes. Two groups of students were matched on five counts, and the same instructor taught both the experimental and the control group. The experiment consisted of a "flexible, group-work teaching plan" in which students engaged in various problem-solving discussions. End of semester results indicated that the experimental group made significant improvement (1950: 34). It is worth noting that Cook is the only sociologist ever to examine empirically the issue of critical thinking with the technical assistance of an educational researcher.

In 1976, Charles Logan attempted to answer his own question, "Do Sociologists Teach Students to Think More Critically?" His search for an answer led to a research project and an experimental class in social problems. Logan measured critical thinking by constructing twenty short-statement items which "contained some example of uncritical or non-scientific thinking about social phenomena" (1976: 32). This test was used to measure both inclination and ability to think critically. Using a quasi-experimental design, he obtained pre- and post-test data on students at all levels of instruction (elementary undergraduate, advanced undergraduate, and graduate). With the exception of his own experimental class in which the skills of critical thinking were stressed, he found low scores in all classes. Logan's conclusions are in direct contrast to prevailing professional myths about the sociology curriculum (1976: 37):

> Neither inclination nor ability to think critically or scientifically about statements on social issues is associated directly or substantially with increasing levels of sociological training in the ordinary course of the curriculum.

Logan explains his negative findings in the following terms (1976: 41):

> The professed concern among sociologists with teaching students to think more rationally and scientifically about social phenomena may be to a considerable degree lip service that masks a hidden curriculum. Sociology professors may in fact be more concerned with teaching students what to think than how to think. That is, they may be more concerned with instilling certain ideas, beliefs, values, and ideologies in their students than with teaching them how to critically and scientifically evaluate those ideas.

This provocative study deserves much attention, but it contains a central flaw: lack of a clear definition of critical thinking. The measuring instrument does not specify those aspects of critical thinking which are under investigation. Testing items did identify various logical fallacies, problems of overgeneralization, and circular reasoning, but these items were not classified in any systematic fashion.

Logan's success with a teaching strategy to promote critical thinking is consistent with empirical studies reviewed in this chapter. Students can be coached to learn successfully specific critical thinking skills found on tests administered at the end of the semester (Glaser, 1941; Lyle, 1958; Oliver & Shaver, 1966; Cook & Cook, 1950; Bloom and Rakow, 1969). These investigations confirm Glaser's assessment of twenty-five studies conducted in the 1920s and 1930s. To quote from his synopsis of the pioneering literature (1941: 69-70):

> There is no evidence that students acquire skill in critical thinking as a necessary by-product of the study of any given subject. On the other hand, almost any subject or project can be so taught as to put pupils on guard against hasty generalization, contradictory assertions, and the uncritical acceptance of authority. In general the research indicates that if the objective is to develop in pupils an attitude of "reasonableness" and regard for the weight of evidence and to develop ability to think critically about controversial problems, then the component attitudes and abilities involved in thinking critically about such problems must be set up as definite goals of instruction.

In the spirit of a "new Columbus," Logan may merely have rediscovered what an educational researcher had clarified from primary sources thirty-five years ago. Nevertheless, Logan still made an important contribution by placing critical thinking back on the sociology teaching and research agenda. To paraphrase Andre Gide, perhaps it was all said before, but since no one was listening, it must be said again.

Discussion of knowledge available about critical thinking has emphasized various conceptual taxonomies and empirical investigations of learning outcomes in the classroom. This review of literature does not exhaust the research topics about critical thinking by any means. Four significant areas of inquiry are not covered. First, not discussed is the thorny issue of the relationship between intelligence (as measured by I.Q. tests) and critical thinking (as measured by critical thinking tests). What is the meaning of a moderately high correlation between these two types of tests? Educators differ in their interpretations (Morse & McCune, 1964: 9 vs. Oliver & Shaver, 1970: 487; see also Rust et al., 1962; Furst, 1950). Second, no mention is made of the relationship between critical thinking abilities and hierarchical stages of human development. In recent years educators have devoted much attention to various typologies of human development (e.g., Piaget, Kohlberg, Loevinger, Erikson). These stages seem to parallel ideas of higher mental processes which place critical thinking near the top of the lofty peaks of cerebral performance. Yet the stages lack precision and the match with critical thinking skills is far from perfect (Ennis, 1975). What can be said about various stages of development and the capacity to succeed with various critical thinking tasks? Third, I have side-stepped the ambiguous literature on critical thinking as an attitude. Clearly, critical thinking is more than proficiency of logical skills; as early as 1910 Dewey recognized the importance of personal attitude in solving problems as well as the mental rigor necessary to carry out the task. Notions of open and closed minds for both teachers and students are worthy of further research commitment (Mahan, 1955; Ehrlich, 1961; Kemp, 1962). Finally, I have not discussed classroom conditions in which critical thinking might be optimized. Does the facilitation of critical thinking necessarily require small classes of homogeneously bright students who enjoy a Socratic dialogue with the teacher? Further inquiry into the relationship between critical thinking processes and group dynamics is in order (Oliver & Shaver, 1966; Rice, 1978; Baker & Jones, 1979). Some knowledge is available on all four topics, but more research is needed to advance the improvement of instruction. I will not explore these topics further. This chapter will continue in another direction by exploring knowledge needed to improve the testing of critical thinking outcomes.

Knowledge Needed to Improve Assessment of Critical Thinking

Three strategies are suggested to improve the sociology teacher's capacity to assess critical thinking. The first approach involves a survey of existing instruments which have been constructed by experts. The second strategy calls for the development of analytical tools which can be used to assess existing banks of sociology test items. And last, knowledge is needed to help teachers construct intelligently their own evaluation instruments. These three strategies are designated respectively as inductive, critical, and constructive. These strategies involve more than the identification or construction of test items. It is of equal importance to understand the purposes for which the tests are intended. The purposes of testing, in turn, are related to the educational goals of instruction. In short, one must always consider the techniques of testing in the broader context of teaching and learning.

Inductive Strategy: Survey of Existing Instruments. During the past half century numerous attempts have been made to construct reliable and valid tests of critical thinking. The Illinois Critical Thinking Project has reviewed 38 tests (Stewart, 1978; Tomko and Ennis, 1979), and I have located additional evaluation instruments. Given the limitations of space, I will mention some of the most important tests in terms of two broad categories: (1) expertly-designed tests which have maximized reasoning skills and minimized the importance of content (see Chart 1); (2) published tests which provide useful formats to assess learning outcomes. The existing content in most of these tests, however, is either obsolete or irrelevant for sociologists (see Chart 2).

According to Tomko and Ennis, the Watson-Glaser Critical Thinking Appraisal and the Cornell Critical Thinking Test are "the only general tests of critical thinking currently available" in the market of expertly-designed tests (1979: 6). Both tests include manuals with discussion of validity and reliability data. The designers of these tests have specified carefully the skills of logical reasoning they attempt to assess. The format is constructed carefully with multiple-choice objective items. Since these tests measure generic reasoning skills, it is difficult to know how potentially relevant

Chart 1

Instruments to Assess Critical Thinking
(Tests Designed by Experts and Currently Available or being Developed)

Author	Name of Instrument	Testing Format of Critical Thinking Instrument	Aspects of Critical Thinking Being Assessed
1. Goodwin Watson & Edward Glaser (1964)	Watson-Glaser Critical Thinking Appraisal (forms Zm and Ym)	Series of relatively short passages with one hundred objective multiple choice items	1. Ability to make inferences on basis of given data; 2. Ability to recognize assumptions; 3. Deduction; ability to decide whether a statement follows necessarily from premises; 4. Interpretation; ability to weigh evidence; 5. Ability to evaluate strength or weakness of arguments
2. Robert Ennis & Jason Millman (1971)	Cornell Critical Thinking Test (level Z)	Series of relatively short passages and a chart from a scientific experiment are used as complex stimuli for 52 objective multiple-choice items	Determine student's proficiency in judging whether: 1. a statement follows from the premises, 2. something is an assumption, 3. an observation statement is reliable, 4. an alleged authority is reliable, 5. a simple generalization is warranted, 6. a hypothesis is warranted, 7. an argument depends upon an ambiguity, 8. a statement is overvague or over specific 9. a reason is relevant (Tomko & Ennis, p. 4)
3. McBer & Company	Test of Thematic Analysis	Open ended essay responses which compare and contrast two groups of stories	"The ability to form and articulate complex, abstract concepts which will order and arrange a confusing array of inchoate data" (1978: 3)
4. McBer & Company	Analysis of Argument	Two open-ended essays responding to a controversial statement; first essay refutes statement while the second essay supports it	"The capacity to think clearly and dispassionately while seeing elements of truth in all sides of a heated controversy" (1978: 3)
5. American College Testing Program	College Outcomes Measures Project (COMP) Measurement Battery	Sixty open-ended questions clustered around 15 activities	Ability to communicate, to solve problems, to clarify values
6. American College Testing Program	COMP - Objective Test	Sixty-three multiple choice questions requiring application of general knowledge and skill to realistic situations	Ability to communicate, to solve problems, to clarify values

BEST COPY AVAILABLE

Chart 2

Instruments to Assess Critical Thinking
(Published Tests Relevent but Not Readily Available for Classroom Use)

Author	Name of Instrument	Testing Format of Critical Thinking Instrument	Aspects of Critical Thinking Being Assessed
7. Paul Baker & Janet Jones (1979) Baker (1979)	Creative Writing on a Social Problem	Forty-five minute essay about a specific social problem (e.g., poverty, crime)	Ability to: (1) define the problem, (2) present supporting evidence, (3) articulate cause-effect relationships, (4) offer a logically consistent statement regarding a solution (or non-solution) to problem
8. Charles Brown (1950)	Name not given	A short passage with 5 objective multiple-choice items to be ranked from 1 (best) to 5 (worst) reasons for a given conclusion	Ability to choose rational evidence to support a given conclusion
9. Paul Dressel & Lewis Mayhew (1954)	A Test of Critical Thinking in the Social Sciences	Passages, graphs and charts presented as complex stimuli; series of objective multiple-choice items follow each passage or chart	Ability to: (1) Identify central issues, (2) recognize underlying assumptions, (3) evaluate evidence or authority, (4) recognize limitations of data, (5) establish relationships, (6) draw warranted conclusions
10. Paul Dressel & Lewis Mayhew (1954)	A Test of Critical Thinking (form G)	Passages and short paragraphs presented as complex stimuli; series of objective multiple-choice items follow each passage	Ability to: (1) define a problem, (2) select pertinent information, (3) recognize assumptions, (4) formulate relevant hypothesis, (5) draw conclusions validly
11. Albert Levi (1948)	Inventory of Social Understanding I	150 short statements about social issues; 1 of 5 responses solicited: (1) Statement which can be proved to be true (2) Statement can be proved false (3) Statement for which there is a preference though it cannot be proved true or false (4) Statement is rejected though it cannot be proved false (5) Statement falls in none of the above four classes	Ability to discriminate between statements of social fact and statements of social value
12. Charles Logan (1976)	Name not given	Open ended responses to 20 short statements	Inclination to recognize and ability to adequately criticize various logical fallacies and non-scientific assertions
13. Horace Morse & George McCune (1964)	Testing of Study Skills and Critical Thinking: Section II	Series of short statements requiring respondent to classify as fact or opinion	Ability to distinguish between facts and opinions or between facts and an author's interpretation of facts
14. Horace Morse & George McCune (1964)	Testing of Study Skills and Critical Thinking: Section III	Series of short statements requiring students to classify as primary sources or secondary accounts	Ability to distinguish between two kinds of statements about evidence: primary sources and secondary accounts
15. Horace Morse & George McCune (1964)	Testing of Study Skills and Critical Thinking: Section V	A group of key references are listed and a series of questions are given; students are asked to match a specific reference with a given question	The ability to find and critically assess relevant information on a given social issue
16. Horace Morse & George McCune (1964)	Testing of Study Skills and Critical Thinking: Section VI "Exercise in Open-Mindedness"	Series of short statements requiring student to respond in terms of All, Most, Many, Some, No	Ability to identify qualified generalizations and thereby avoid unqualified assertions of all or none

BEST COPY AVAILABLE

Chart 2 (continued)

17. Horace Morse & George McCune (1964)	Testing of Study Skills and Critical Thinking: Section XIII "Exercise in Consistency"	A paragraph is presented and it is followed by a series of questions; student classified each question as consistent or inconsistent with ideas found in the paragraph	Ability to recognize questions which are pertinent to a particular line of reasoning
18. Horace Morse & George McCune (1964)	Testing of Study Skills and Critical Thinking: Section XV	Series of short statements requiring students to classify each statement as thoughtful or emotional	Ability to distinguish between biased and objective statements
19. Horace Morse & George McCune (1964)	Testing of Study Skills and Critical Thinking: Section XVI	A generalization is presented; it is followed by a series of statements to be classified as supporting the generalization or irrelevant to the generalization	Ability to recognize empirical statements which support generalizations. "Given a general statement and a body of data, what are the data which support the general statement?" (1964: 80)
20. Horace Morse & George McCune (1964)	Testing of Study Skills and Critical Thinking: Section XVII	A short statement is presented; it is followed by a series of questions to be classified as meaningful or not meaningful	Ability to recognize pertinent questions which extend empirical knowledge about a given statement
21. Donald Oliver & James Shaver (1966)	Social Issues Analysis Test I (SIAT #1)	An argumentative dialogue is presented and followed by a series of objective multiple-choice items	"Assess how well students can identify selected intellectual operations occurring in an argumentative dialogue" (1966: 191). Intellectual operations are defined as: (1) Problem identification and differentiation; public controversy requires differentiating: (a) empirical or factual controversy, (b) value controversy, (c) definitional controversy; (2) Making explicit important assumptions, (3) Identifying and using appropriate strategies for dealing with different types of problems, (4) Identifying and using common dialectical operations: (a) generalizing, (b) specifying, (c) qualifying, (5) Identifying relevant problems
22. Donald Oliver & James Shaver (1966)	Social Issues Analysis Test 2 (SIAT #2)	An argumentative dialogue is presented and followed by a series of objective multiple-choice items	"Assess both how well the student can identify the substance of an argumentative dialogue (in contrast to the intellectual operations occurring in it) and the student's ability to select the best rebuttals which might be used to counter statements made in the dialogue" (1966: 194)
23. Donald Oliver & James Shaver (1966)	Social Issues Analysis Test 3 (SIAT #3)	Interview format. Interviewer reads aloud a dialogue, while student follows on his own copy. Student is asked open-ended questions about types of problems and appropriate strategies for solution	"Requires the student to differentiate and identify various types of disagreements as well as suggest now the disagreements might best be clarified or resolved through different analytical strategies" (1966: 204)
24. Donald Oliver & James Shaver (1966)	Social Issues Analysis Test 4 (SIAT #4)	Free discussion of 12 students who respond to a controversial case. Discussion is scored for each individual participant who expresses one or more "items of thought"	Two scoring systems are superimposed on each other: (1) static categories--(a) value judgments, (b) legal claims, (c) factual claims, (d) definitional claims, (e) relevance, (f) debate strategy; (2) dynamic categories--(a) consistency-inconsistency, (b) specification-generalization, (c) qualification

BEST COPY AVAILABLE

Chart 2 (continued)

25. David Simon (1977)	Name not given	Students select various magazine articles by American social critics and score the articles according to the following ideological orientations: (1) pluralist, (2) radical liberal, (3) conflict, (4) other	Ability to analyze journalistic articles according to problem loci, problem solution, and problem outcome; this analytical model provides the means to determine ideological categories of pluralist, radical liberal, and conflict
26. Eugene Smith Ralph Tyler & Evaluation of the Eight Year Study (1942)	Interpretation of Data Test	Passages, charts, tables, graphs presented as complex stimulus; series of statements follow and respondent judges on the basis of data provided, whether the statement is: (1) true (2) probably true (3) not sufficient to indicate whether there is any degree of truth or falsity (4) probably false (5) false	Test explores two related questions: (1) "To what extent does the student recognize the limitations of the data?" "How accurately does the student perceive various types of relationships in the data?" (1942: 56-58)

BEST COPY AVAILABLE

they are to sociology. Oliver and Shaver see little value in such pencil-and-paper tests for occasions of civic education (1966: 184), but this negative opinion needs to be further examined empirically by sociologists in their classrooms.

In recent years considerable attention has been given to issues of academic accountability and the educational value of a liberal arts education. One aspect of this soul searching is a renewed effort to measure the lofty claims of general education. Responding to the need to know more about learning outcomes, two testing firms are developing new critical thinking tests. Since these efforts are in the developmental stage, they will be discussed as tentative projects with a future yet to be written in the annals of the testing enterprise.

McBer and Company is developing a Comprehensive Cognitive Assessment Battery (CCAB) which probes those basic competencies essential for sophisticated adult roles. They are examining dimensions of higher education which lead students "to become critical and discerning thinkers, competent problem solvers, and persons who will take the initiative in finding solutions to problems" (1978: 1). CCAB consists of four tests (Test of Thematic Analysis, Analysis of Argument, Learning Style Inventory, and Self-Definition), but the first two instruments more directly concern issues of critical thinking and sociology. The Test of Thematic Analysis is described as "a measure of critical thinking ability." "This test assesses the process of forming, articulating, and using concepts when comparing, contrasting, and rearranging information in order to draw conclusions" (1978: 2; see also Whitla, 1977: 9). Analysis of Argument measures intellectual flexibility. "This test reflects the capacity to think clearly and dispassionately when dealing with viewpoints that involve controversial and emotional issues" (1978: 2). Designed to be "content free," both tests require students to respond in writing to various stories and controversial statements.

In collaboration with twelve colleges and universities, The American College Testing Program (ACT) recently launched a major ten-year project to develop new methods of assessing intended learning outcomes of general education. The College Outcome Measures Project (COMP) evaluates six types of cognitive activities: three content-related domains (Functioning within Social Institutions, Using Science and Technology, and, Using the Arts) and three process-related domains (Communicating, Solving Problems, and, Clarifying Values). The three content-related domains are matched with the three process-related domains creating a basic three by three matrix (Forrest, 1978: 3-8). ACT also makes use of various "real life" materials drawn from the print and electronic media. The first phase of test development includes (1) a Measurement Battery, (2) an Objective Test, and (3) an Activity Inventory. The Measurement Battery consists of several tests in which students respond to various stimulus items by written essays or audiotaped statements. "The Objective Test is the product of an effort to create proxy measures for all components of the Measurement Battery" (Forrest, 1978: 15). Finally, The Activity Inventory uses a multiple choice format to assess the quality and quantity of participation in out-of-class activities related to the six outcome areas (Forrest, 1978: 4).

McBer and ACT seem more concerned with the pragmatic claims of critical thinking than some earlier designers of critical thinking tests. Their instruments are more sensitive to criterion-referenced testing which seeks to relate critical thinking to successful adult role performances. This new interest in competency-based assessment may appear sloppy, but nonetheless offers promise of a new breakthrough in testing strategies.[4]

Sociology teachers need reliable knowledge about the critical thinking capacities of their students. The expertly-designed tests mentioned above provide some of the empirical tools essential to begin such a task. But these available tests do not exhaust the range of empirical instruments which might be used to assess critical thinking. Chart 2 identifies several tests also relevant to teachers of sociology. In many cases the original instruments use content items which are obsolete or of minimal value to sociology. But the testing formats and suggested strategies of inquiry are highly relevant. Sociologists can benefit from these earlier efforts as they revise and adapt instruments to their own specific questions.

The 26 tests of critical thinking mentioned in Charts 1 and 2 are highly diverse in form, degree of complexity, range of topics, and function. I will review briefly some of the salient aspects of this diversity.

Three types of testing formats are used. The first type is an open-ended essay form in which students are required to articulate their best thoughts with minimal direction from the test itself (Tests 3, 4, 5, 7, 12, 23). Clearly the most extreme open-ended test reported in this review of literature is Oliver and Shaver's Social Issues Analysis Test #4 in which discussion groups graple with a controversial issue (Test 24). The second type of format is the presentation of paragraphs, charts, tables, or graphs as complex stimuli; a series of structured questions (usually multiple-choice items) follow each complex stimulus (Tests 1, 2, 6, 8, 9, 10, 15, 17, 19, 20, 21, 22, 26). Simon's test (Test 25) is both open-ended and prestructured. Students are asked to analyze a wide variety of magazine articles with a particular classification scheme. But this scheme is not a totally closed system and on some occasions, students may reject the adequacy of the scheme altogether. The third type of format is the least complicated for students, but sometimes frustratingly complex for test analysts (Levi, 1948). A series of one-sentence statements requires the student to respond according to a predetermined classification scheme (Tests 11, 13, 14, 16, 18). Expected judgments are often dichotomized (e.g., a statement is either fact or opinion), but occasionally responses may be more elaborate (e.g., Test 11).

Earlier discussion of various definitions of critical thinking revealed numerous complex components in the term. It is not surprising to rediscover this complexity when one examines the many aspects of critical thinking being assessed by the 26 tests. The diverse topics cannot be classified easily by a typology of mutually exclusive and exhaustive categories. The tests are designed with various degrees of analytical rigor, and in some cases information is limited to a brief discussion and a few sample items in a research article. These limitations do not preclude a few general comments about various aspects of critical thinking being scrutinized by the test makers. Some tests are comprehensive measures of critical thinking (Tests 1, 2, 5, 6, 7, 9, 10, 21, 22, 23, 24); other tests cover a more narrow range of topics (Tests 3, 4, 8, 11, 13, 14, 15, 16, 17, 18, 19, 20, 25, 26). The most comprehensive tests are the Cornell Critical Thinking Tests; Ennis and Millman attempt to cover more [illegible] aspects of critical thinking in one 50-minute test than any other instrument in the literature. At the other extreme Morse and McCune provide an extensive battery of single-focused tests of critical thinking. Charts 1 and 2 also provide some indication of the topics most frequently covered by tests of critical thinking. Several tests attempt to measure the ability to make inferences from a given body of data, find assumptions, develop a consistent line of reasoning, define problems, and clarify value judgments. It seems highly probable that virtually all of the topics mentioned in the 26 tests are found in Ennis's recent comprehensive statement of rational thinking (1979).

Finally, a few comments are in order regarding the functions of the 26 tests. In some instances researchers have used their instruments for purposes of norm-referenced testing (Tests 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 21, 26). But these same instruments are capable also of being used for purposes of criterion-referenced testing. Another group of tests is intended exclusively for criterion-referenced testing purposes (Tests 7, 13, 14, 15, 16, 17, 18, 19, 20, 25). Several tests have been used at one time or another as diagnostic instruments (Tests 1, 2, 7, 11, 13, 14, 15, 16, 17, 18, 19, 20, 26). Baker has constructed a test for formative evaluation (Test 7) and many others are capable of use in this fashion. Finally, several tests are intended for summative evaluation (Tests 1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 21, 22, 23, 24, 26). Researchers have used several tests as pre-post test devices to measure the effectiveness of various instructional innovations (Tests 1, 9, 10, 12, 21, 26). On the other hand, many tests are not sufficiently rigorous at this point for such research purposes. Given this diversity of formats, topics, and functions, sociology teachers must weigh thoughtfully the facets of critical thinking which concern them and the purposes of assessment which guide their inquiry.

Critical Strategy: Evaluating Existing Test Items. Charts 1 and 2 do not exhaust the battery of tests available to sociology teachers. With monotonous regularity the mail brings new textbooks with accompanying test banks of quiz items. What do sociologists know about items in these test banks? Do these test items measure various aspects of critical thinking? What conceptual guidelines are needed to assess the critical thinking capacities of these items? Such questions deserve much attention. But educational researchers have not developed analytical schemes or operational techniques to assess systematically the existing test items. One of the few scholarly works to explore this problem is Richard Anderson's (1972) pioneering work on assessing tests of comprehension. Using Bloom's taxonomy as a point of departure, Anderson appraises various techniques of pencil-and-paper tests essential to measure the first two levels of knowing: knowledge and comprehension. He points out that verbatim questions (even when words have been rearranged) are never capable of measuring comprehension. Assessing comprehension requires such methods as paraphrasing questions, substituting particular terms for superordinate terms, substituting general terms for specific terms, identifying new instances which discriminate from non-instances. Using the scheme outlined by Anderson, I examined 100 items from Ian Robertson's test bank. Robertson has authored the leading text on the market today (1977a), but on issues of critical thinking it is a dismal failure. The vast majority of items are verbatim; some items are misleading, and few used the range of testing possibilities outlined by Anderson. Robertson's test bank did not test students at the second rung on Bloom's ladder of higher mental processes (1977b).

Teachers of sociology need to develop a theoretical framework to assess the adequacy of critical thinking instruments. Guidelines are needed to evaluate existing sociology test items as to their usefulness in revealing capacity for critical thinking. Everett K. Wilson raises some helpful questions for such an endeavor (1979, personal correspondence). First, does the test item get at students' ability to take things apart, to see the elements (analytic skill)? Second, does the item measure students' ability to see connections, systemic properties (synthetic skill)? Third, does the test item get at students' ability to extend beyond the instant case or situation? That is, does it test skills in detecting analytically isomorphic populations or situations? Wilson's three questions are a valuable beginning; the analytical work of Ennis and others could be used to extend the list.

Constructive Strategy: Guidelines for Teacher-Made Tests. I have now reviewed various tests which claim to assess critical thinking. I have also suggested that sociologists need to develop an evaluation system to determine the extent to which commercial test banks assess critical thinking. A third strategy is suggested to improve the instruction of sociology. It is the need to develop theoretical criteria and operational techniques to construct "teacher-made tests" of critical thinking. I know of very little literature, except Tomko & Ennis (1979) which deals with this topic directly, but numerous books and articles have been written about teacher-made tests (McKeachie, 1969; Green, 1963; Grunland, 1968; Berg, 1965). This "how to . . . " literature is not grounded in careful research; furthermore, it perpetuates a widely-accepted dichotomy among educators. It seems that both amateur teachers and testing experts insist on dividing all tests into two mutually exclusive categories: essay tests and objective tests.

Conventional discussion of teacher-made testing often seems to boil down to a basic dilemma. Essay tests are easy to construct, but hard to score. Objective tests, on the other hand, are hard to construct but easy to score.

In McKeachie's words, "It appears to be true that examinations which are easiest to construct are the most difficult to grade and vice versa" (1968: 127). Unfortunately, this mentality pervades the practical advice to teachers who need to construct their own tests. This is a false dichotomy which fails to acknowledge that the construction of meaningful tests is always difficult regardless of the formats being used. Likewise, the generation of scores never escapes the thorny question of validity. Given the administrative purposes of testing in formal educational settings and the constraints of time and resources for testing any given group of students, problems of determining a meaningful sample of worthwhile questions will always be on the teacher's desk. Teachers sample a small segment of a person's life experience and attempt to generalize to the outer limits of their data. Teachers' work is complicated further by attempts to measure the "higher mental processes" of critical thinking. This is the practical agenda of all teachers who design their tests. The issues are not clarified by either/or decisions about essay versus objective exams.

Knowledge needed to improve teacher-made testing must begin with the teacher's instructional objectives. If sociology teachers make claims of facilitating critical thinking in their classrooms, then they must devise some system of assessing learning outcomes commensurate with these claims. They also must determine what aspects of critical thinking they are promoting. Some aspects of critical thinking are more adaptive to objective exams than others. For example, any teacher who is interested in the topic of inference can find many working models in the current collection of published critical thinking tests. Inference questions have two basic statements: one statement takes the form of a conclusion and the other takes the form of supporting evidence. Any number of objective testing possibilities can be designed around these two components. On the other hand, critical thinking tests which attempt to relate complex theories to novel problem-solving tasks are probably best assessed by essay questions.

Deliberate consideration of instructional objectives often leads to a second basic issue. Sociology teachers claim frequently that knowledge from textbooks and classroom lectures is relevant and transferable to non-textbook literature and to settings outside the classroom. In solemn voice, the clichés come easily to sociology teachers on the first day of class: "This textbook and my lectures will help you better understand events reported in the daily newspaper." If instructional objectives intend to provide useful learning opportunities for life experiences beyond the classroom, then testing devices should extend to those life experiences deemed worthy of the benefits of formal education. If sociologists believe that their instruction will benefit all persons who read newspapers, then testing devices are needed to assess the student's newly-acquired capacities to better understand social knowledge found in newspapers. Such tests can be either objective or essay. For example, Robert Ennis has designed an objective critical thinking test which uses a simulated newspaper speciman (Moorburg Letter Test); I have designed several essay tests which make extensive use of journalistic sources (Baker, 1979). Testing learning outcomes which stretch beyond the immediate boundaries of textbook or lecture requires careful selection (or imaginative creation) of "relevant" case materials (Newmann & Oliver, 1967).

A case can be made against objective testing; Hugh Petrie has made the best case I have ever read (1979). But his line of reasoning does not lead necessarily to an exclusive use of essay tests. There are alternative possibilities of using both objective and essay formats in deliberate tension with each other. I have been using such a hybrid arrangement for several years. My rationale and testing device are rather pragmatic and straightforward. Pre-structured answers leave no room for inquiry and often create the misleading notion that successful academic work is little more than agreeing with the Answer Man. Fools consistently sucker for false answers while the wise student agrees with the test maker's choice. But somewhere between foolishness and wisdom there is a gray zone of intelligent doubt about specified options. According to Ennis, critical thinking is the ability to assess the adequacy of statements; it seems reasonable to extend Ennis's point by suggesting that the mark of a critically-thinking student is the capacity to assess the adequacy of the teacher's question with its accompanying options. Perhaps for logically sound reasons, none of the options fit the question, or perhaps two options can be defended as logically sound choices.

Rather than insisting on an either/or position, I have developed a multiple-choice system which permits an open-ended response to the entire question/answer complex. After careful reflection on the question and the options, students may choose to offer a short written reaction which takes exception to the prearranged claim of measuring some principle of sound reasoning. In some cases these written responses may demonstrate incompetence on the part of the student; on other occasions the written comments call attention to the inadequacy of the test maker's line of reasoning or choice of words. In Dewey's terms, students were confronted with an obstacle, and they were given the opportunity to intelligently overcome it.

Developing new teacher-made tests which probe issues of critical thinking may require a radical reassessment of conventional teaching/learning expectations. Sociology teachers may need to re-examine some of the basic premises of their instruction. Hopefully, such reassessment will lead to meaningful innovation. I am currently developing an alternative teaching/learning system for basic courses in sociology (Baker, 1975; Baker, 1979). One aspect of this work is the creation of several critical thinking tests. These tests are designed to assess two fundamental cognitive skills: (1) the ability to assess the adequacy of others' statements about social reality, (2) the ability to create logically sound statements about social reality. Mastery of these generic skills allows students to go beyond the immediate text literature of sociology. I require students to examine sociology critically by comparing its truth claims with those of common sense and journalism.

Students are never asked to accept sociological knowledge as a worthy end in its own right. They are never tested on their ability to memorize or comprehend conventional textbook knowledge. As rational thinkers students must scrutinize sociological ideas with the same critical tools used for all other domains of social thought. Sociology is not immune from critical assessment in the sociology classroom. In short, critical thinking begins when professional knowledge is denied special privileges of passive receptivity.

Conclusion

I would like to suggest a simple syllogism. Since assessment of learning outcomes should coincide with instructional objectives, and since the instructional objectives of many (if not most) sociologists include a verbal commitment to critical thinking, it seems reasonable to conclude that conscientious teachers would be ever mindful of the task of evaluating critical thinking in the classroom. Yet such is not the case. Sociologists talk a good line, but they come up short at delivery time. They seem satisfied with teaching/learning arrangements which demand little more of students than the replication of favorite concepts, famous names, and club-house quarrels among the elite.

Critical thinking is one of those buzz words which seems to mean just about anything to anybody. It comes in handy for rhetorical justification of sociology as a curricular enterprise, but the term rarely is examined as an analytical concept capable of empirical scrutiny. When one reviews the conceptual literature, it is clear that the best researchers and educators have been probing a complex set of logical operations and interdependent mental processes. Different aspects of critical thinking have been measured successfully and these measuring instruments have become crucial dependent variables in numerous educational experiments. Numerous educators have demonstrated repeatedly that instructional arrangements can be designed to facilitate the learning of critical thinking. These successful learning outcomes have been measured for the short run, but little is known about the long-run educational processes that facilitate critical thinking as life-long intellectual orientation. Much work remains to be done. It will begin when teachers of sociology become conscious of the gap between what they say and what they do.

Appendix A

Addresses of Critical Thinking Tests Currently Available

Cornell Critical Thinking Test, Level Z
Robert Ennis
Rational Thinking Project
Education Building
University of Illinois
Urbana, Illinois 61801

Watson-Glaser Critical Thinking Appraisal
Harcourt, Brace and Jovannovich, Inc.
New York, New York

College Outcome Measures Project (COMP)
Audrey Forrest, Director
The American College Testing Program
P.O. Box 168
Iowa City, Iowa 52240

Comprehensive Cognitive Assessment Battery (CCAB)
McBer and Company
137 Newbury Street
Boston, Massachusetts 02116

Notes

*I am grateful to Robert Ennis, Mark Felson, and Everett K. Wilson for critical comments to an earlier draft of this chapter. Research for this chapter was supported by a grant from Illinois State University and the Fund for the Improvement of Post-Secondary Education, HEW.

1. In a provocative paper, Gregory Bateson identifies three levels of learning, and suggests that this hierarchy is part of the larger scheme of all human communication (1972: 279-308).

2. This point is worth emphasizing. Critical thinking is not a residual category which includes all intellectual abilities above the level of rote memorization.

3. I say rarely because Max Black includes elementary symbolic logic in his influential book, Critical Thinking (1952).

4. Since most sociology teachers may be unaware of these expertly-designed tests of critical thinking, Appendix A provides addresses for further inquiry.

References

ANDERSON, R. A. (1972) "How to construct achievement tests to assess comprehension." Review of Educational Research 42: 145-170.

ANGELL, R. (1964) Reasoning and Logic. New York: Appleton-Century-Crofts.

AUSUBEL, D. (1961) "Reception versus discovery learning in classroom instruction." Educational Theory 11: 21-24.

BAKER, P. (1975) "Social awareness project." Teaching Sociology 3: 74-80.

--- (1979) "Dialectical teaching: an instructor's manual to social awareness modules." (unpublished)

BAKER, P. and J. JONES (1979) "A case study of benign disruption in the classroom." Teaching Sociology, forthcoming.

BATESON, G. (1972) Steps to an Ecology of Mind. New York: Ballantine Books.

BENDIX, R. (1970) Embattled Reason: Essays on Social Knowledge. New York: Oxford University Press.

BERG, H. /ed./ (1965) Evaluation in Social Studies, Thirty-Fifth Yearbook of the National Council for The Social Studies. Washington, D.C.: National Council for the Social Studies.

BEYER, B. (1971) Inquiry in the Social Studies Classroom. Columbus: Charles E. Merrill Publishing Company.

BIGGE, M. (1976) Learning Theories for Teachers. New York: Harper & Row Publishers.

BLACK, M. (1952) Critical Thinking. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

BLOOM, B., HASTINGS, J. and G. MADAUS (1971) Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill Book Co.

BLOOM, B. and E. RAKOW (1969) "Higher mental processes," pp. 594-601 in R. Ebel (ed.) Encyclopedia of Educational Research. New York: The MacMillan Co.

BRADSHAW, T. and S. MCPHERRON (1978) Data Report for Issues and Resources in the Undergraduate Sociology Curriculum (ASA Projects on Teaching Undergraduate Sociology). Washington, D.C.: ASA Teaching Resources Center.

BROWN, C. (1950) "An experimental diagnosis of thinking on controversial issues." Speech Monographs 17: 370-377.

CHICKERING, A. (1976) "Developmental change as a major outcome," pp. 62-107 in M. Keeton (ed.) Experiential Learning. San Francisco: Jossey-Bass Publishers.

COOK, A. and E. COOK (1950) A Sociological Approach to Education. New York: McGraw-Hill Co.

COOK, A. and R. KOENINGER (1939) "Measuring learning outcomes in introductory sociology." Journal of Educational Sociology 13: 208-225.

DENZIN, N. (1970) The Research Act. Chicago: Aldine Publishing Co.

DEWEY, J. (1933) How We Think. Lexington, Mass.: D. C. Heath and Co.

DOOB, L. (1950) "Propagandist vs. propagandees," pp. 439-458 in A. Gouldner (ed.) Studies in Leadership. New York: Harper & Brothers.

DRESSEL, P. and L. MAYHEW (1954) General Education: Explorations in Evaluation. Washington, D.C.: American Council on Education.

EBEL, R. (1978) "The case for norm-referenced measurements." Educational Researcher 7: 3-5.

ENNIS, R. (1962) "A concept of critical thinking." Harvard Educational Review 32: 81-111.

--- (1969) Logic in Teaching. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

--- (1964) "Operational definitions." American Educational Research Journal 1: 183-201.

ENNIS, R. and J. MILLMEN (1971) Manual for Cornell Critical Thinking Test, Levels X and Z. Urbana, Illinois: University of Illinois.

ENNIS, R. (1973) "On causality." Educational Research 2: 4-11.

--- (1975) "Children's ability to handle Piaget's propositional logic: a conceptual critique." Review of Educational Research 45: 1-41.

--- (1979) "Logic, rational thinking, and education." Philosophy of Education 1979: Proceedings of the Thirty-Fifth Annual Meeting of the Philosophy of Education Society (ed.) Jerrod Coombs, Philosophy of Education Society.

FORREST, A. (1977) Comp: College Outcome Measures Project. Iowa City: American College Testing Program.

--- (1978) Comp: Annual Report 1978. Iowa City: American College Testing Program.

FURST E. (1950) "Relationship between tests of intelligence and tests of critical thinking and of knowledge." Journal of Educational Research 43: 614-25.

GARDNER, W. (1970) "The current scene," pp. 1-10 in W. Gardner and F. Johnson (eds.) Social Studies in the Secondary Schools. Boston: Allyn and Bacon, Inc.

GIROUX, H. (1978) "Writing and critical thinking in the social studies." Curriculum Inquiry 8: 291-310.

GLASER, E. (1941) An Experiment in the Development of Critical Thinking. New York: Teachers College, Columbia University.

GLASER, R. (1963) "Instructional technology and the measurement of learning outcomes." American Psychologists 18: 519-521.

GREEN, J. (1963) Teacher-Made Tests. New York: Harper & Row, Publishers.

GRONLUND, N. (1968) Constructing Achievement Tests. Englewood, N.J.: Prentice-Hall, Inc.

HAMKELTON, R., SWAMINATHAN, H., ALGINA, J. and D. COULSON (1978) "Criterion-referenced testing and measurement: a review of technical issues and developments." Review of Education Research 48: 1-47.

HAAS, J. (1977) The Era of the New Social Studies. Boulder, Colorado: Social Science Education Consortium.

LEE, A. and E. LEE (1939) The Fine Art of Propaganda. New York: Harcourt, Brace and Company.

LEVI, A. (1948) General Education in the Social Studies. Washington, D.C.: American Council of Education.

LOGAN, C. (1976) "Do sociologists teach students to think more critically?" Teaching Sociology 4: 29-48.

MAHAN, T. (1955) "Human judgment: can the classroom improve it?" Journal of Educational Research 49: 161-169.

MCBER and CO. (1978) Users Guide for the Comprehensive Cognitive Assessment Battery (CCAB). Boston: McBer and Company.

MCKEACHIE, W. (1969) Teaching Tips: A Guide for the Beginning College Teacher. Lexington, Mass.: D. C. Heath and Company.

MORSE, T. and G. MCCUNE (1964) Selected Items for the Testing of Study Skills and Critical Thinking. Washington, D.C.: National Council of the Social Studies.

NEWMANN, F. and D. OLIVER (1967) "Case study approaches in social studies." Social Education 31: 108-113.

OLIVER, D. and J. SHAVER (1966) Teaching Public Issues in the High School. Boston: Houghton Mifflin Company.

OSBORN, W. (1939) "An experiment in teaching resistance to propaganda." Journal of Experimental Education 8: 1-17.

PETRIE, H. (1979) "Against 'objective' tests: a note on the epistemology underlying current testing dogma." (unpublished)

POPHAM, W. (1978) "The case for criterion-referenced measurements." Educational Researcher 7: 6-10.

RICE, T. (1978) "Cognitive mapping: small groups technique in teaching." Teaching Sociology 5: 259-274.

ROBERTSON, I. (1977a) Sociology. New York: Worth Publishers, Inc.

ROBERTSON, I. (1977b) Test Bank to Accompany Robertson Sociology. New York: Worth Publishers, Inc.

RUSSELL, D. (1960) "Higher mental processes," pp. 645-661 in C. Harris (ed.) Encyclopedia of Educational Research 3rd Edition. New York: The MacMillan Company.

RUST, V., JONES, R. and H. KAISER (1962) "A factor analytic study of critical thinking." Journal of Educational Research

SCRIVEN, M. (1967) "Methodology of evaluation," pp. 40-66 in R. Tyler, R. Gagil and M. Scriven (eds.) Perspectives of Curriculum Evaluation. Chicago: Rand McNally & Co.

SIMON, D. (1977) "The ideologies of American social critics." Journal of Communication 3: 40-50.

SJOBERG, G. (1973) Instructor's Guide in Cities: Where People Live and Why. Boston: Allyn and Bacon.

SMITH, E. and R. TYLER (1942) Appraising and Recording Student Progress. New York: Harper and Brothers.

STEWART, B. (1978) Reviews of Selected Tests in Critical Thinking Test File. Urbana, Illinois: Illinois Critical Thinking Project.

TOMKO, T. and R. ENNIS (1979) "Evaluation of informal logic competence." (unpublished)

TURNER, S. (1979) "The concept of face validity." Quality and Quantity 13: 85-90.

TYLER, R. (1936) "The relation between recall and higher mental processes," pp. 6-17 in C. Judd (ed.) Education as the Cultivation of the Higher Mental Processes. New York: MacMillan.

UNIVERSITY OF ILLINOIS (1979) College of Arts and Sciences "Final report on the requirement in general education."

WHITLA, D. (1977) Value Added: Measuring Outcomes of Undergraduate Education. Cambridge, Mass.: Office of Instructional Research and Evaluation, Harvard University.

WILSON, E. (1970) "Inductive methods in teaching sociology," pp. 113-120 in M. Krug, J. Poster and W. Gillies (eds.) The New Social Studies. Itasca, Illinois: F. E. Peacock Publishers, Inc.

--- (1979) Personal letter to author (July 1979).

Designs and Procedures for Evaluating Course Innovations

James W. Michaels

Virginia Polytechnic Institute and State University

The publication of the journal *Teaching Sociology* since late 1973 has apparently done much to increase the salience of teaching concerns in general, and course innovations in particular. The importance of the journal as a dissemination outlet is reflected in a recent survey of the 1977 issues of 59 social science and higher education journals by Goldsmid and Davis (1978). The authors found *Teaching Sociology* to account for one-third (25 of 75) of all articles deemed relevant to teaching sociology, and four-fifths (25 of 31) of the relevant articles published in sociology journals. Although related disciplines have similar journals, this represents the highest concentration of articles on the teaching of a social science in a single journal.

The authors note that, although most of the relevant articles report course innovations, empirical evaluations of the effects of innovations are rare. This is particularly true of the course innovations reported in *Teaching Sociology*.[1] Although most authors do at least make casual comments about the presumed benefits of course innovations, such subjectively based testimonials hardly constitute sufficient evidence of relative efficacy. It would indeed be ironic if social scientists who demand hard evidence of effects in non-applied investigations were less demanding in applied cases in which students' and instructors' outcomes may be at stake. Thus, although the potential impact of *Teaching Sociology* may be considerable, the actual impact may be attenuated by uncertainty regarding the demonstrable effects of innovations reported.

Both the problem identified and the solution proposed here have been considered by others (e.g., Campbell et al., 1977). The proposed solution calls for the application of the same criteria and standards typically applied to evaluative experimental research in general. Unfortunately, most sociologists have had little training in experimental designs and procedures. The purpose of this chapter is to consider feasible experimental designs and procedures for the systematic evaluation of course innovations. In contrast to the excellent general explication on experimental designs in educational research by Campbell and Stanley (1963), this chapter treats specific problems and procedures. Although operating on different levels of abstraction, our shared concern with strengthening internal validity should be evident.

Experimentation and Internal Validity

Campbell and Stanley (1963: 2) regard experimentation as the only means for settling questions about educational practice, the only means for verifying education improvements, and the only means for establishing a tradition in which improvements can be introduced without faddish disregard of old wisdom in favor of inferior novelties. The major attraction to experimentation is that it provides the opportunity (not guarantee) to strengthen internal validity. Internal validity refers to the certainty with which one can conclude that a particular independent variable (e.g., a course innovation), rather than extraneous or unidentified variables, produced an observed effect on a particular dependent variable (e.g., learning outcomes). The major advantage of experimentation is that it provides the opportunity to control extraneous variables (usually by constancy) that might otherwise be alternative explanations for observed effects.

This chapter is based on the premise that systematic evaluations of course innovations can approximate the standards applied to experimental research in general. As in any experimental research, all extraneous variables cannot be controlled. However, the fact that some exceptions to the ideal may be unavoidable is not sufficient justification for disregarding experimentation and getting really sloppy.

Feasible Designs and Associated Problems

Intergroup Designs

One of the more feasible true experimental designs considered by Campbell and Stanley (1963: 25-27) is the Posttest-Only Control Group Design. This design requires two groups, treatment (e.g., innovative procedures) and control (e.g., non-innovative or traditional procedures), with measures of dependent variables following (or during) treatments. Thus, two sections of the same course taught by a single instructor may serve as treatment and control groups.[2]

Unfortunately, this design, as with all true experimental designs, involves a procedure that is frequently impractical with college classes--random assignment of subjects to treatments (i.e., students to sections in this case). When random assignment is impractical, the quasi-experimental Non-equivalent Control Group Design (Campbell and Stanley, 1963: 47-50) may be adopted and appropriate pre-treatment measures used to determine the extent of initial differences between sections on variables likely to be related to eventual learning outcomes. Pre-treatment measures that might be relevant include grade point average, college major, year in colleges, number of (and grades in) previous sociology courses, and even pretests on substantive material.

Fortunately, with sufficiently large sections, initial differences between sections are likely to be minor.

Class size and schedule differences create additional problems of control. Although small-to-moderate class size differences may be only a minor concern, the effects of class size differences may depend on absolute differences in enrollment. For example, a difference of 20 students may be less of a threat if the classes are composed of 35 and 55 students respectively than if they are composed of 15 and 35 students. Class size differences may also affect some dependent variables (e.g., class participation) more than others (e.g., student mastery of material). Schedule differences (i.e., time of day, days of week, and length of class meetings) should also be minimized. In particular, sections that meet on different days (which usually means they will also differ on number and length of class meetings) and early morning and late afternoon classes should be avoided.

On the basis of pre-treatment measures and class size and schedule differences, one section may appear to have an initial advantage in probable course outcomes. In keeping with the rule that the innovation should never be given an advantage that is not part of the innovation itself, the less advantaged section should be assigned to the innovation treatment. The problem of determining with certainty which is the advantaged section can be avoided by adopting the alternative design described below.

Intragroup or ABAB Designs

This design (called the Equivalent Time-Samples Design by Campbell and Stanley, 1965: avoids all the noted problems of the previous designs by using a single section as both treatment and control in temporal rotation. This design is most appropriate, however, when the treatment is expected to have a powerful, but transient or reversible, effect. This is because it is necessary to demonstrate at least partial return to control (baseline) levels of performance following each replication of treatment. Thus, as desirable as an effective non-reversible treatment might be otherwise, the effectiveness of such a treatment cannot be scientifically demonstrated using this design. An additional limitation of this design is that statistical tests of significance are problematic because the usual independence assumptions are not met. For this reason it is important to show (usually in graphic form) substantial treatment effects (e.g., performance gains) with each replication of treatment conditions and partial return to baseline (e.g., performance losses) with each replication of control conditions.

Application of the ABAB design to course innovations raises additional considerations. First, frequent measures (e.g., once a week and on the same day of the week) of dependent variables are highly desirable. Second, each replication should be of sufficient duration for dependent variable measures to reach a moderately stable state. In this regard a partial (ABA) design may be used under the quarter system or whenever longer treatment duration is deemed more crucial than the gain in internal validity. Another possible adaptation is a combination of inter- and intragroup designs involving only one application of each treatment to each of two sections. Thus, treatments may be applied to one section in AB sequence and to the other in BA sequence (see Lamberth and Knight, 1974). Third, several changes in course procedures during the term may be somewhat awkward. Finally, such changes may increase student awareness of experimentation and produce reactive effects.

Intergroup designs dominate the empirically evaluated course innovations reported in the teaching journals of the social sciences. Although this domination is likely to continue, intragroup designs may represent a viable alternative in some cases. One such case is where there is only one section of the course, as is usually the case with graduate and upper level undergraduate courses. As might be expected, the vast majority of empirically evaluated course innovations reported involve lower level courses and large classes. Consequently we know little about the relative efficacy of alternative course procedures for upper level and smaller classes. In such cases an intergroup design may be less appropriate because it would require the comparison of two classes separated by a time lapse sufficient to constitute a threat to internal validity.

Independent Variables

Independent variables in this case are those that constitute the innovation and are of two types--instructional and non-instructional. Instructional variables involve the materials and methods of instruction, whereas non-instructional variables involve course policies and procedures other than instructional.

Instructional Variables

This category includes all instructional materials (e.g., textbooks, outside readings, class handouts) as well as all instructional techniques applied in and out of class (e.g., lecture, discussion, gaming and simulation, audio-visual aids, student presentations, readings, and other assignments).

Very few studies have tested the relative efficacy of different reading materials. Yet courses for which there is proliferation of textbooks covering essentially the same material (e.g., introductory sociology) provide an excellent opportunity to test the relative efficacy of different texts in getting material across to students. Several introductory texts also offer coordinated student study guides or workbook supplements whose effects could be examined (e.g., supplement used versus supplement not used).

The greater number of studies varying class format may be partially accounted for by our insistent faith that what is done in class has a major impact on learning outcomes. In the past, studies comparing lecture, discussion, and other class formats have been particularly popular (e.g., see the review by Bligh, 1974). Although inconclusive in terms of general efficacy, such studies have identified several possible interaction effects, e.g., the relative efficacy of stance, students' familiarity with the topic, class size, instructor preference, and, most importantly, how each is specifically implemented. Obviously the lecture format provides the instructor greater opportunity to communicate information regarded as important, whereas the discussion format provides greater opportunity for student participation.

Several currently popular techniques appear to represent attempts to provide greater opportunities for experiential learning. Such techniques include gaming and simulation (e.g., SIM-SOC, MOD-SIM, CLUG), field trips, and various small gropp or class projects including "Doin Soc" (Clifton, 1973, 1976).

In testing the relative efficacy of instructional procedures two related problems occur with sufficient frequency to warrant a cautionary note. The first problem is the failure to fully implement the different treatments. For example, a lecture format in which lecture constitutes only three-fourths of class time, and a discussion format in which the instructor does most of the talking, may not constitute sufficiently different treatments. Because unintentional bootlegging of aspects of one treatment into another may attenuate effects, it is highly desirable to demonstrate empirically that different treatments have, in fact, been implemented. In the case of lecture and discussion formats, for example, an observer could classify and count communications by the instructor and students, and these frequencies could be reported in tabular form. A related problem identified by Goldsmid and Davis (1978) is that all too frequently the instructional formats involved are not described in sufficient detail for replication by others. Actually, this appears to be somewhat of a one-sided problem in that the innovative format is usually described in detail, whereas the control format is sometimes described by little more than a label. The solution to both problems is to attend as carefully to the operationalization, measurement, and description of independent variables as to dependent variables.

Non-Instructional Variables

This category includes what is to be learned, how and when learning is to be demonstrated, and performance standards and grade contingencies. Although non-instructional variables are often ignored as a basis for strengthening academic learning and performance, recent evidence indicates a renewed interest in such variables. In fact, such variables are considered to be essential components of several successful instructional systems emphasizing high mastery of material and/or self-pacing (e.g., PSI).[3]

Most such systems include the specification of material to be mastered (i.e., instructional objectives) and how mastery is to be demonstrated; self-paced learning and frequent opportunities to demonstrate mastery; and criterion referenced, rather than peer referenced, grading. Thus, the primary concern is to design a system that provides students greater incentives and opportunities to demonstrate mastery. Because such systems generally require more time to implement and are more difficult to coordinate than more traditional procedures, it may be tempting to adopt only those components easiest to implement. However, recent research designed to test the independent contributions of the separate components of PSI suggests that it may be the unique combination of components that is primarily responsible for the effectiveness of such systems (Calhoun, 1976).[4]

Nevertheless, continued research on the independent effects of such instructional variables as the number, timing, and types of tests, retests, and exams; test and exam options or exemptions; performance standards and grade contingencies; and available options may identify effective procedures that are relatively easy for a single instructor to implement.

Dependent Variables

Dependent variables in this case are those representing outcomes, consequences, or effects of course innovations. Perhaps the dependent variable of greatest concern is student learning or mastery of material as evidenced by test scores throughout the term. Although the tests for two or more sections need not be identical, sections should be compared only on identical test items. It may also be desirable to include different types of test items in the comparisons in order to test for possible treatment-by-item type interactions effects. In addition to tests, mastery and application of material may be reflected in student presentations, term papers, and research projects. However, any item or application that may involve a significant subjective component should be evaluated in the blind.

A second important type of dependent variable that is often ignored as a supplementary measure is student self-reports and evaluations. Student self-reports may be important as supplementary measures for several reasons. First, they may provide evidence regarding the extent to which treatments were implemented. Second, they may provide supplementary or collaborative measures for some, and the only measures for other, dependent variables. Finally, student self-reports may provide helpful suggestions for future application.

Questionnaire items used should be designed with these considerations in mind, for most student course evaluation questionnaires will be practically useless for these purposes. Student self-reports should be regarded with caution, however, and should not be used a primary data on learning outcomes. A study by Samuels (1977), for example, suggests that innovations which are relatively ineffective for strengthening such variables as class participation, class attendance, and test scores may nevertheless be perceived by students as being more valuable than less innovative class procedures.

Other measures that may be considered either as intervening or as dependent variables include class attendance, class participation, peer tutoring, and consultations with the instructor. Generally, however, such social process variables should probably be regarded as intervening variables which may or may not affect student learning outcomes.

A final consideration often omitted is instructor outcomes. Course innovations are frequently time consuming, and their effects are not always as great as expected. Considering the time constraints on both instructors and students, a cost-benefit analysis might be both informative and useful. Such an analysis would relate the extra time involved to the gains in student learning. Although there are no doubt other more subjective and personal outcomes more difficult to measure, these may not be shared by others considering adoption. The ultimate measure might be whether the innovation is permanently adopted were it not for the fact that inferior innovations may continue to be used as long as they satisfy the personal needs of the instructor.

Reporting Innovations and Effects

Presumably most investigators will not be content to complete a systematic evaluation of an innovation simply to satisfy their own curiosity--particularly if the innovation is successful. Reports of innovations and their effects should satisfy the criteria applied to research reports in general.

It would be helpful to begin the report with a statement of the problem addressed and citations of the relevant literature supporting the contention that the innovation should be superior to alternative procedures for alleviating the problem. Although most effective course innovations operationalize one or more established principles of learning related to course design, communication, motivation or performance, according to the review by Goldsmid and Davis (1978), authors frequently fail to acknowledge such ties. Second, the innovation, control, and measurement procedures should be described in sufficient detail in the methods section to allow replication by others. Summary data (e.g., means and the standard deviations) and the results of statistical tests for effects on all dependent variables should be reported clearly in the results section. Although simple tests for differences in means should be sufficient in most cases, additional factors may be considered with analysis of variance and, with treatments represented as dummy variables, even multiple regression. In some cases it may be helpful to present data in graphic form--particularly when an ABAB design has been used. Finally, it would be helpful to others considering adoption to present an estimate of the cost-benefit ratio in terms of extra hours per unit of gain.

A conservative posture is advised when considering generalizability or external validity of observed effects. Considering the frequency of higher order interaction effects in educational research, the question of generalizability is best answered through replication by different instructors in different courses composed of different types of students.

Conclusion

This chapter has called for more systematic evaluations of course innovations reported in social science teaching journals--particularly Teaching Sociology. Specifically, there is little reason why evaluations of course innovations cannot approximate the criteria and standards applied to experimental evaluation in general. Considering most sociologists' lack of training for, and predisposition toward, such experimental research, this chapter has considered several feasible experimental designs, associated problems, and remedial procedures for systematically evaluating the effects of course innovations. Several possible independent and dependent variables were also identified. The purpose of going to the extra trouble is not simply to mix research with instructional concerns, but to strengthen our confidence that a particular innovation was or was not effective. One would hope that this might reduce the frequency with which we disregard old wisdom in favor of inferior novelties.

Notes

1. My review of course innovations reported in Teaching Sociology and Teaching of Psychology from 1974 to 1978 found 7 of 41 (17%) to report empirical evaluations in the former journal, whereas the corresponding figures for the latter were 20 of 44 (45%). The empirical evaluations included experimental and control groups, quantitative measures of dependent variables, and statistical tests for effects.

2. The traditional labels "experimental group" and "control group" may be somewhat misleading in that both groups are controlled and constitute comparison groups in the experiment. The label "treatment group" is also misleading because it suggests that only that group receives treatment. Nevertheless the traditional labels are retained here with the understanding that both or all groups compared have received some sort of treatment.

3. Although PSI (Personalized System of Instruction) is perhaps the best known of these, Bassis and Allen (1976, 1977) have described and tested a computer based system called TIPS (Teaching Information Processing System) that may be more efficient for large classes.

4. For tips on evaluating such "modularized" systems of instruction see Fitzgerald, 1977.

References

BASSIS, M. and J. ALLEN (1976) "Individualization and economy for mass instruction." Teaching Sociology 3: 185-190.

--- (1977) "Teaching introductory sociology with TIPS: an evaluation." Teaching Sociology 4: 141-154.

BLIGH, D. (1972) What's the Use of Lectures? Baltimore: Penguin.

CALHOUN, J. (1976) "The combination of elements in the Personalized System of Instruction." Teaching of Psychology 3: 73-76.

CAMPBELL, D. and J. STANLEY (1963) Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally.

CAMPBELL, F., KAUFMAN, D., GIECHRIST, C., HANCOCK, R., KLOBAS, J., and A. SUNDGREN (1977) "Experimenting with curricular designs: some problems and possibilities." Teaching Sociology 5: 49-64.

CLIFTON, A. K. (1973) Doin Soc: An Undergraduate Field Manual. Champaign, Illinois: Stipes.

--- (1976) "Doin soc: an evaluation of actively involving introductory students in the work of sociology." Teaching Sociology 3: 138-147.

FITZGERALD, J. (1977) "On recognizing innovation: a primer to evaluate 'new and improved' teaching methods." Teaching Sociology 5: 87-99.

GOLDSMID, C. and E. DAVIS (1978) "The teaching of sociology, 1977: annotated bibliography and review." Teaching Sociology 5: 325-360.

LAMBERTH, J. and J. KNIGHT (1974) "An embarrassment of riches: effectively teaching and motivating large introductory psychology sections." Teaching of Psychology 1: 16-20.

SAMUELS, F. (1977) "Actors, skits, and concepts." Teaching Sociology 4: 155-166.

Assessing Teacher Effectiveness

Theodore C. Wagenaar
Miami University

The assessment of teacher effectiveness has a long history of empirical inquiry, particularly in the areas of educational measurement and psychology. The topic has received increasing attention in recent years due to demands from both sides of the professor's desk: administrators and students. Administrators have been increasingly concerned with services rendered in their attempt to attract and retain students and in response to legislative demands for accountability. Similarly, students have become increasingly vocal in their demands for effective teaching, reflecting the general trend toward increased client demands for accountability. As reflections of these trends, Kelly and Auleta (1978) note that fewer than 50 percent of educational institutions employed student ratings in the mid-sixties while over 85 percent now employ these ratings, and Landers (1978) cites a recent decision by Idaho's State Board of Education mandating the use of student evaluations in tenure and promotion decisions.

This chapter will summarize what is currently known about teacher assessment in higher education. Comments will then be made about aspects of teacher assessment in need of further research.

A Few Basic Assumptions

I shall begin with a few basic assumptions commonly shared by researchers in the field of teacher assessment. First, teaching is a phenomenon which can be investigated. The elements of teaching can be identified and researched and related to various other variables. This is not to say that there is one best way to teach; neither is it to say that teaching is not subject to various philosophical discussions on the meaning of teaching. It is only to say that, like other phenomena, teaching in higher education is analyzable and, therefore, amenable to scientific investigation.

Second, teaching quality varies from high to low effectiveness. Again, researchers are not in a position to prescribe a single model of teaching effectiveness. Effectiveness may vary by discipline, course level, nature of students, and characteristics of the instructor. But I am assuming that some of these dimensions can be identified and empirically investigated.

An example of one such dimension is student learning; certainly effective teaching should result in higher amounts and levels of learning than less effective teaching.

Third, I assume that social science methodology can be applied to teacher effectiveness, particularly strategies for measurement. Undoubtedly, measures of teacher effectiveness share the reliability and validity problems common to most social scientific inquiry. But at least ordinal level measures can be developed for several of the dimensions.

Fourth, teaching effectiveness should be viewed contextually. That is, effective teaching is not a product solely of the individual's traits. It also is influenced by the social ecology of the classroom, the organizational structure of the department, and the socio-political organizational climate of the entire college or university.

Fifth, I am assuming that multiple methodologies and sources of data are better than a single methodology or source. Hence, research incorporating observation design, survey design, and documentary analysis is preferred over research employing only one design. Similarly, studies utilizing data from colleagues, administrators, alumni, the instructor, and students are preferable to those utilizing only one of these data sources.

The Definition of Effective Teaching

Before considering the various approaches to assessing teacher effectiveness, attention should be paid to the key variable: effective teaching. Unfortunately, researchers do not often consider this measurement issue; discussions of teaching competence are far more frequently found in the non-research based exhortations about college teaching. Since this article is a review of knowledge available and knowledge still needed regarding teaching effectiveness, I will not consider the more pedagogical and philosophical approaches to the issue.

However, a few points are important to keep in mind as one reviews the research in this area. First, many definitions of teaching and effective teaching exist. This factor alone makes comparability across studies somewhat difficult. Second, few researchers deal with the conceptualization and operationalization of the concept teacher effectiveness in the same way they deal with other concepts. In fact, most researchers simply accept the operationalizations of existing teacher evaluation forms.

Third, many pedagogical and philosophical discussions of teacher effectiveness emphasize the changes brought about in the student, while most researchers emphasize the techniques of teaching and tend to ignore the products of teaching. Fourth, little attention has been paid to constructing theories of teaching; without some theoretical base, research on teacher effectiveness often results in a rather descriptive account of interrelationships among variables.

Fifth, it is important to note that when outcome variables are considered, the most common measure is the end-of-term examination. However, this measure of student learning is subject to criticism on the basis of both reliability and validity. For example, many if not most end-of-term examinations are not comprehensive but cover only the last segment of a course. Also, the testing strategies employed by many instructors say little for the validity of such measures; most tend to use hastily prepared multiple-choice type questions. Furthermore, a variety of extraneous variables intrude on such final exam measures, including student overload, maturation effects, and instructor overload. Hence the use of end-of-term exams as a definitive measure of student learning seems questionable.

Having considered a few preliminaries, I will now review some of the empirical literature of teacher effectiveness. Since over 90 percent of the research employs student ratings, the bulk of this section of the chapter will review these studies. But I will first briefly review the research on colleague evaluation. After considering both these data sources, I will conclude the paper with some reflections on knowledge still needed.

Research Based on Colleagues

In terms of evaluation for promotion and tenure, the opinions of colleagues are paramount. Undoubtedly these opinions play a larger role than do student evaluations. It is probably safe to conclude that virtually all educational institutions employ colleague evaluation for personnel decisions. However, fewer than 5 percent actually based these evaluations on classroom visits (Selden, 1976).

There is very little research on peer evaluation. In the only review of the literature on colleague evaluation, Batista (1976) laments the paucity of articles on peer evaluation, the even greater paucity of research articles, the lack of adequate conceptualization regarding the role of peers in the total evaluation process, and the lack of systematization in peer evaluation procedures. The majority of the articles reviewed are policy oriented and tend to be polemical (i.e. why peer evaluation should or should not occur, the role of peers vs. administrators).

The few research articles which do examine peer evaluation tend to focus on comparing peer assessments with students' and administrators' assessments. For example, Centra (1973) noted that peers generally rate instructors more highly than do students (4.5 vs 4.0 on 5 point scale), and that the reliability coefficient for colleague evaluations is lower than that for students (.57 vs. .85). Perhaps these results reflect the possibility that colleagues employ different criteria than do students, particularly such non-teaching criteria as research. However, it should be noted that the correlations between students' and colleagues' evaluations are consistently moderately positive, which attests to at least a moderate amount of validity for student evaluations (Ballard, Rearden, and Nelson, 1976).

A much lower level of agreement exists between peer and administrative evaluations, particularly above the chair level. This discrepancy is due to the fact that different sets of criteria are employed (Meany and Ruetz, 1972). Also interesting to note is the conclusion that administrators do not employ a consistent objective method in their evaluative efforts (Meany and Ruetz, 1972). With regard to faculty opinion about administrative competence to evaluate teaching, Sheehan (1975a) notes that only eight percent of the faculty he studied felt that administrators were qualified to evaluate their teaching (corresponding figures for students, colleagues, and self are 86 percent, 75 percent, and 66 percent respectively).

Hence we see that little peer evaluation exists, even though faculty think that their peers (and their students) are qualified to evaluate their teaching. Perhaps one reason for the lack of peer evaluation is the relative absence of established criteria and systematic observation guides. Comprehensive guides such as those developed by Helling (1976) may engender both increased use of peer evaluation and increased amounts of research in this area.

Research Based on Students

Research on student evaluation of teaching has a long and sometimes controversial history. This emphasis on student evaluations in the assessment literature may simply reflect the fact that data generated by such evaluations are easily quantified and readily available. Also, more data have become available as student evaluations have become more popular.

Given this immense data source and the fact that much of the research is *post hoc* in nature, many studies have been completed which fail to follow some of the basic rules for quality research. As a result, one can find at least some research to support virtually any position on the topic. The research summarized below strives for consistency by focusing on major findings in quality studies.

The validity of student evaluations was noted above in the discussion noting positive correlations between faculty and student evaluation. Students and faculty also agree on the key ingredients of an "ideal" professor. And ratings by students correlate at least moderately positively with ratings by measurement specialists (Landers, 1978).

The reliability of student ratings has been well documented. No matter what the measure of reliability employed, coefficients generally exceed .90. Stability is further documented by the high correlation between end of term evaluations and 3-10 year subsequent evaluations. Hence, student ratings are highly reliable and at least moderately valid (Costin, Greenough, and Menges, 1971; Doyle, 1975).

A major focus has been on the relationship of various teaching behaviors (as rated by students) to student learning (as usually measured by final examinations). Various correlations have been reported, ranging from -.80 to +.80. However, the median correlation is minimally positive. This range of results can partially be explained by noting that a variety of factors other than the instructor's teaching can contribute to final exam or course grade; Tuckman (1975) notes that student ability and desire to learn are far more predictive of learning than are instructor characteristics. Also, most exams are not valid indications of the behaviors measured by student ratings (Grasha, 1977). Furthermore, if one assumes that the student is more responsible for learning than anything the instructor does, the inconsistent results are not surprising. In sum, the inconsistent results reflect how little we know about the linkages between specific aspects of teaching and learning.

Student ratings can also be related to affective outcomes. Research in this area suggests that a mildly positive relationship exists between ratings and students' decision to continue work in an area. Also, highly rated instructors are more frequently selected for future courses.

Student demographic characteristics (e.g., age, sex, level) are generally unrelated to ratings, suggesting that ratings are largely independent of such characteristics. Neither are ratings of teacher skill related to students' personality traits or to their attraction to teachers as persons (Grush and Costin, 1975).

Regarding faculty characteristics, most demographic factors are minimally related to student ratings (sex, rank, experience). Neither are self-assessed personality characteristics of faculty related to student ratings; student-assessed personality characteristics of faculty are somewhat related to student ratings (Landers, 1978). Instructors' knowledge of subject matter does not correlate with student ratings, suggesting either that most instructors are knowledgeable in their fields or that students are not capable of making valid assessments of the trait.

An area of major dispute among faculty is the relationship of research productivity to teacher effectiveness; arguments exist to support both a positive and a negative relationship. However, the research quite clearly indicates a negligible relationship between the two factors, suggesting that competent researchers are neither better nor worse in the classroom than professors with lower research productivity.

Another area of dispute among faculty involves the case of the popular teacher. The research does not support the conclusion that the easy, humor-oriented popular teacher scores significantly higher on student ratings. Neither does it appear to be the case that students attempt to damage an instructor through ratings (Grasha, 1977).

We have examined student and faculty characteristics separately. There is some support for the conclusion that faculty and student characteristics interact in their effects. Most notably, the higher the similarity between student and instructor personality characteristics, the higher the ratings by students. Grasha (1977) suggests that a match between student learning style and instructor teaching style yields higher ratings and increased learning. Similarly, instructors focusing on bright students receive higher ratings from these students, and instructors focusing on poorer students receive higher ratings from these students. Hence, a congruency model is suggested for maximal teacher effectiveness.

Several course characteristics are also related to student ratings. Instructors teaching smaller classes and those teaching elective courses receive higher ratings than their counterparts. Also, instructors of upper level courses receive higher ratings than instructors of introductory level courses. It should be noted, however, that some interaction effects occur since most upper level courses tend to be small and elective; although all three variables have independent effects, the course level seems to be the most salient (Landers, 1978). Regarding discipline, Linsky and Strauss (1973) conclude that sociology faculty receive lower ratings than several other disciplines; other researchers indicate that few significant differences exist among the disciplines (Dwyer, 1968).

Administration effects are also present. Instructors receive higher ratings when they tell their students that the results will be used for decision making purposes. Instructors also receive higher ratings if they remain in the room while students fill out evaluation forms.

Some research has also been done on the effects of information provided by student evaluations on instructors. More change in subsequent teaching behaviors occurs when the results are discussed with colleagues. Also, very highly and very poorly rated instructors are less likely to change instructional techniques than moderately rated instructors; highly rated instructors may feel no need to change while poorly rated instructors may feel that change may not be worth the effort (although poorly rated instructors confronted repeatedly with negative feedback do make some positive changes in their teaching). More change occurs when the instructor gives self a better rating than do students. However, it should be noted that most instructors do not use the results for modifying their teaching (Landers, 1978).

Which factors emerge as the most salient in the various student rating forms? On the basis of reviews by Hildebrand and Wilson (1970), Kulik and Kulik (1974) and others, five general factors emerge. One is skill-communication and includes such components as use of examples, pacing, knowledge of student understanding, good answers to questions, and clarity. A second is structure-organization and includes such components as clear objectives, clear expectations, careful grading, synchronization of text with class, and preparation. A third is enthusiasm-rapport and includes such components as tolerance, friendliness, interaction with students, humor, and enthusiasm for subject. A fourth is subject knowledge and includes such components as the use of many sources, many examples, lecturing without notes, recent citations, and discussion of new ideas. A fifth is difficulty-workload and includes such components as stress, high expectations, many assignments, and constant probing. Parenthetically, Sheehan (1975b) reports that this last factor is one of few significantly related to student achievement.

Most student rating forms also include a global item on general teaching effectiveness. My own research on over 7,000 students in over 150 courses indicates that the two most predictive factors of this overall rating are skill-communication and enthusiasm-rapport; the contributions to the total variance explained (which is 71 percent) are 57 percent and eight percent respectively, and the beta values are .36 and .20 respectively. Also relevant is the fact that most items on student rating forms are at least moderately positively related.

Given the popularity of student evaluations, a few problems with their use should be noted. First, few forms include valid measures of student learning, making inferences about the results of varying levels of teacher competence difficult. Hence we know little about which characteristics of teachers and teaching contribute to critical thinking abilities, knowledge of principles and theories, retention of facts, the ability to learn independently, and other outcomes for students (for an exception to this conclusion, see Johnson, Rhodes, and Rumery, 1975b). Similarly, we know little about how students have changed over the term, which differs from a simple end-of-term measure of learning since students may have entered a course with previous knowledge.

Second, only a few forms employ comparative data. Given the high average scores most instructors receive on most forms (students are basically quite kind), the use of comparative data is mandatory. Such comparative data should be broken down by class size, discipline, status of course (required or elective), institutional type, and other relevant variables.

Third, the measurement and evaluation uses of student ratings should be distinguished. Student evaluations do provide a measure of the presence of various teaching behaviors and can therefore be used for diagnostic purposes. But to use them for evaluation purposes requires attention to the goals of the teacher, department, and institution, to the purposes of the evaluation (e.g., salary, promotion), to other sources of data (peer evaluations, course materials), to a clarification of relevant values employed, and to more stringent criteria for reliability and validity measures.

Fourth, results of student evaluations require the availability of qualified personnel to both interpret the results and assist the faculty member in modifying subsequent teaching behaviors. Without such support services, the results of student ratings will remain largely unused. Such support personnel will be useful for developing self-assessment skills among faculty members.

Fifth, attention should be paid to tailoring student evaluations to specific courses and instructors so that utilization and relevance are maximized. Such tailoring does not obviate the use of standardized instruments; it simply yields additional information which may be particularly useful for a given instructor. Sixth, more attention needs to be paid to separating course characteristics from instructor characteristics in both research and interpretation.

Knowledge Still Needed

This review of the literature suggests several weaknesses in the current knowledge about teacher assessment. Knowledge still needed can be categorized into three sections: theory, research design, and unexamined (or underexamined) variables.

Most research on teacher assessment is atheoretical. That is, most research focuses on interrelating a few variables or on statistical analyses of factors of effective teaching as indicated in student rating forms. Little attention is devoted to such prior questions as what is teaching? what is effective teaching? what are the outcomes of teaching? what are the inputs to teaching? what is learning? how do students learn?, and the like. In all the studies I reviewed, only Johnson, Rhodes, and Rumery (1975a, 1975b) consider theory. After reviewing the major approaches (learning outcomes, teacher characteristics, and pedagogical behavior), they develop a normative theory focusing on the development of a capacity for critical norm-oriented human action as a central goal in higher education. In this theory they distinguish teaching from other activities incorporating the same techniques, and they develop a "gnometectonic" sequence integrating pedagogical and student activities. Hence evaluation should focus on descriptions of teacher intent and both student perceptions of these intents and student response to these intents. Their venture is an ambitious one, but it does reflect possible theoretical developments.

In addition to the necessity for attention to pedagogical theoretical development, there exists a necessity for sociological theoretical development. A variety of sociological theories could be used to examine and assess teaching. For example, ethnomethodological approaches could be used to determine the "taken-for-granteds" present in the daily interactions of both instructors and students. Building on Goffman's work, Karp and Yoels (1976) suggest that instructors treat the classroom as an instance of "focused" interaction while students define the classroom more as an "unfocused" situation. Such an approach may clarify our understanding of the role of the instructor and, hence, of teaching.

Additionally, the sociometric theories in sociology may be relevant for understanding classroom ecology, and hence teaching. For example Becker, et al. (1973) examined class size and seating arrangements on student participation. Since student participation may be seen as a fundamental component of learning, such sociometric theories become relevant for understanding the teaching-learning process.

Role theory may also be of use in conceptualizing the various roles instructors fill and the factors associated with successful role performance. Complex organizational theory may be of relevance for understanding the effects of the larger institutional setting on performance at the sub-organizational unit to the classroom. Symbolic interactionism may be of relevance for examining the communication and interaction processes in the classroom, the symbols used in communication, and the subjective meanings of social phenomena for participants. The small groups literature may be relevant for developing hypotheses about classroom interactions. Conflict theory may assist in conceptualizing the often adversary relationships that exist between faculty and students. Structural-functionalism may assist in outlining functional and dysfunctional consequences of current structural arrangements in higher education. And systems theory may be relevant for analyzing the interdependencies among student characteristics, course characteristics, instructor characteristics, and organizational characteristics. To date, few researchers have applied these theories to their investigations.

The second area of further knowledge needed regards research design. We need more studies that are of higher quality. Too many of the studies on teacher effectiveness employ small, isolated samples, lack adequate control or comparison groups, and employ statistics inappropriate for the data. A fundamental methodological issue is the measurement of teaching effectiveness.

Increased attention needs to be paid to developing valid measurements of this concept which are consonant with sound theoretical rationales. Special attention should be paid to the role of criterion-referenced testing. Such developments are particularly necessary in the area of student outcomes, especially in terms of measuring outcomes other than low level cognitive outcomes (see Baker, Chapter 14 in this volume). In short, we need more sound research on the teaching-learning process.

The third area of knowledge needed lies in the consideration of additional variables. Although we know a great deal about the relationships of measured effectiveness to various faculty and student demographic characteristics, we know less about variables more directly related to the teaching-learning process. For example, we know little about the effect of such entering student characteristics as previous knowledge, study habits, and motivational sets. Similarly, we know little about such faculty characteristics as approach to one's discipline (is there a relationship between humanism or any other theoretical orientation and teaching effectiveness?), the experiences of the instructor as a student, reasons for entering college teaching, exposure to the pedagogical literature, teaching styles, and goals, attitudes toward students, nonclassroom interactions, and educational philosophy. We also know little about the interaction effects between faculty and student characteristics (beyond the congruency effect noted above).

We know even less about contextual effects. Are such departmental characteristics as size, tenure ratio, sex ratio, level of politics, extent of standardization in text selection and the like related to teaching effectiveness? Do students in large schools experience different levels of teaching competence from students in small schools? Is the level of decentralization related to teaching styles and teaching competence? Do ties with a larger socio-political environment (such as a legislative) have any significant effects? Does teaching competence vary by institutional type (e.g., four-year vs. two-year institutions)?

More attention also needs to be paid to the role of the above-noted variables and others as control variables in examining basic relationships. For example, research on teacher characteristics and teaching behaviors often fails to control for alternative sources of learning. For example, Leicht and Rumery (1973) note that the effect of reading course materials was double the effect of exposure to lectures. Yet few studies consider alternative sources of learning, much less control for them. Similarly, personality factors may be confounded with teaching behaviors and should therefore be controlled. Other candidates for control variables include textbook characteristics, type of examination, frequency of examinations, temporal point at which evaluation is done, amount of work assigned, frequency of class meetings, and the like. These and other variables noted previously (e.g., experience, class size, instructor sex) should be controlled to validly examine relationships between teaching behaviors and ratings of effectiveness.

Conclusion

In sum, we already know a great deal about teaching and its assessment. But most of this knowledge reflects easily measured variables (e.g., sex, level, experience) that are frequently ancillary to the actual teaching-learning process occurring in the classroom and elsewhere. Current knowledge also reflects a tendency to ascribe meaning to measures rather than measures to meaning. Hence teaching effectiveness is commonly defined as being what the forms measure.

The three major avenues of further work in this area include: (1) theoretical reformulation, (2) application of sociological theories, concepts, and principles, and (3) increased methodological rigor (both in terms of measurement and design). Upon completion of work in these areas, we will be better able to specify the components and consequences of effective teaching, and to make recommendations for assessing teaching effectiveness.

BALLARD, M., J. REARDEN, and L. NELSON (1976) "Students and peer rating of faculty." Teaching of Psychology 3 (April): 88-90.

BATISTA, E. E. (1976) "The place of colleague evaluation in the appraisal of college teaching: a review of the literature." Research in Higher Education 4: 257-271.

BECKER, F. D., R. SOMMER, J. BEE, and B. OXLEY (1973) "College classroom ecology." Sociometry 36 (December): 514-525.

CENTRA, J. A. (1973) "Self-ratings of college teachers: a comparison with student ratings." Journal of Educational Measurement 10: 287-295.

COSTIN, F., W. T. GREENOUGH, and R. S. MENGES (1971) "Student ratings of college teaching: reliability, validity, and usefulness." Review of Educational Research 41: 511-535.

DOYLE, K. O., JR. (1975) Student Evaluation of Instruction. Lexington: D. C. Heath.

DWYER, F. M. (1968) "A review of characteristics and relationships of selected criteria for evaluating teacher effectiveness." University Division of Instructional Services. University Park: The Pennsylvania State University.

GRASHA, A. F. (1977) Assessing and Developing Faculty Performance: Principles and Models. Cincinnati: Communication and Education Associates.

GRUSH, J. E. and F. COSTIN (1975) "The student as consumer of the teaching process." American Educational Research Journal 12 (Winter): 55-66.

HELLING, B. (1976) "Looking for good teaching: a guide to peer observation." Danforth Faculty Fellowship Project Report.

HILDEBRAND, M. and R. C. WILSON (1970) "Effective university teaching and its evaluation," in K. Eble (ed.) The Recognition and Evaluation of Teaching. Salt Lake City: Project to Improve College Teaching.

JOHNSON, H. C., JR., D. M. RHODES, and R. E. RUMERY (1975a) "The assessment of teaching in higher education: a critical retrospect." Higher Education 4 (June): 173-199.

--- (1975b) "The assessment of teaching in higher education: a proposal." Higher Education 4 (August): 273-303.

KARP, D. A. and W. C. YOELS (1976) "The college classroom: some observations on the meanings of student participation." Sociology and Social Research 60 (July): 421-439.

KELLY, E. F. and E. M. AULETA (1973) "The evaluation of post-secondary teaching." Paper presented at annual meeting of the American Association of Colleges of Teacher Education, Chicago.

KULIK, J. A. and C. C. KULIK (1974) "Student ratings of instruction." Teaching of Psychology 1 (December): 51-57.

LANDERS, A. D. (1978) "Evaluating Teachers," pp. 355-375 in O. Milton (ed.) On College Teaching. San Francisco: Jossey-Bass.

LEICHT, K. L. and R. E. RUMERY (1973) Role of Teacher Structuring and Student Structuring of Learning Materials in Student Learning. Washington, D.C.: Department of Health, Education and Welfare, Office of Education, Bureau of Research.

LINSKY, A. S. and M. A. STRAUS (1973) "Student evaluations of teaching: a comparison of sociology with other disciplines." Teaching Sociology 1 (October): 103-118.

MEANY, I. O. and F. J. RUETZ (1972) "A probe into faculty evaluation." The Educational Record 53: 300-307.

SELDIN, P. (1976) How Colleges Evaluate Professors. Croton-on-Hudson: Blythe-Pennington.

SHEEHAN, D. S. (1975a) "Faculty attitudes toward evaluation and teaching improvement." Improving College and University Teaching, Yearbook· 74-77.

--- (1975b) "On the invalidity of student personnel decisions." Journal of Higher Education 46 (November/December): 687-700.

TUCKMAN, H. P. (1975) "Teacher effectiveness and students' performance." The Journal of Economic Education 7 (Fall): 34-39.

Teacher Effectiveness: Some Preliminary Questions

Michael A. Malec
Boston College

In assessing the "Knowledge Available, Knowledge Needed" on teacher and teaching effectiveness, there are three basic questions that should be addressed before we get to three larger questions. The first questions to be faced are:

1. What do we _want_ to know?
2. _Why_ do we want to know it?
3. How shall we _measure_ it?

Then we can ask the questions of main concern:

4. What _do_ we know?
5. What do we yet _need_ to know?
6. How can we use our knowledge to _change_ teachers?

First, _what do we want to know_? Social Psychological theory suggests that most people have a need to evaluate their opinions and abilities. Most teachers, and those who are concerned with teaching, make at least implicit judgments about the abilities of teachers (including themselves). Thus, we want to know: Is Teacher X average (whatever that means), above average, or below average? If we can answer that question, it follows that we would want to know: What is it that makes Teacher X average or above or below? Is it something about X's personality? Voice control? Teaching style or method? Eye contact? Rank? Size of class? Some other clearly identified factor? Some combination of factors?

Second, _why do we want to know it_? Most teacher effectiveness literature assigns two reasons. One is faculty development: we simply want to take steps to insure that X is the best teacher possible. It may be true that great teachers are born, not made. But it is also probably true that poor teachers can be made into good teachers, and that good teachers can be made better. The faculty development model is prompted both by the benign wish of professionals to do a good job and by the more recent demands of education consumers for the accounting of their tuition dollars. A second reason for knowing something about teacher effectiveness is _faculty evaluation_ or, more crassly, reward and punishment. The rewards an institution can offer a faculty member--promotion, tenure, salary increments, and the like--are usually based in the university's statutes on the criteria of scholarship, teaching, and service, but many faculty feel that it is only scholarship, as measured by publications, which really counts. Various constituencies are coming to feel that the importance accorded teaching in determining institutional rewards should be increased; to do this requires some judgment of teaching effectiveness.

Third, how shall we measure it? In many respects this is the most vexing question. As social scientists, we are especially concerned with our methods. We know that measurement is crucial. Our journals abound with debate on the general problems of reliability and validity and how to achieve them. We have great difficulty in getting representatives of various schools of thought to agree on the measurement of social class or anomie or any other concept. We have equal difficulty in persuading our faculty colleagues that this procedure or that method can assess their teaching effectiveness.

There have been three traditional ways of measuring teacher effectiveness. First, and most frequent, is the use of student responses to surveys. These range from underground critiques to formalized, institutionally sanctioned questionnaires which are computer-scored and result in a mass of quantified material, perhaps augmented by some qualitative commentary. A second form of assessment is based on some type of peer or superior review. Colleagues or a Chairperson or some other official might, for example, visit a classroom one time or several and offer some form of judgment. Finally, there is self-evaluation, a process in which, ideally, a teacher reflects upon past semesters, compares achievement with previously stated objectives, determines what changes, if any, should be made, and sets goals for the future.

Gardner has provided a thorough review of five evaluation schemes which extend the above three-factor model. He argues "that because the selection of an evaluation methodology is so often determined by latent political or convenience factors, many evaluations are destined to fail before they begin" (Gardner, 1977: 573). His five approaches to evaluation are summarized in Figures 1 and 2 and are briefly described.

1. Evaluation as Professional Judgment. Here, the assumption is that the best evaluation is the expert opinion of a qualified professional. We use this model of evaluation in, for example, doctoral examination committees or when journal editors act as referees in selecting manuscripts for

	Professional Judgment	Measurement	Congruence Between Performance and Objectives	Decision Oriented	Goal-Free/ Responsive
Principle Focus	Expert opinion of qualified professional (s).	Measurement of results, effects, or performance, using some type of formal instrument (test), questionnaire, etc.	Comparison of performance or product with previously stated standards of performance, goals or objectives.	"Delineating, obtaining and providing useful information for judging decision alternatives."	Identification and judgment of actual outcomes (irrespective of goals, standards, etc.) and/ or the "concerns of constituents."
Examples	Accreditation teams. Doctoral Committees. Peer review of grant proposals. Referees for selection of manuscripts for publication. Promotion/Tenure decisions.	GRE scores. Faculty activity questionnaires. Attitude surveys. Teaching effectiveness questionnaires.	Teacher certification based on achievement of prescribed competencies. Evaluation of academic departments on the basis of stated goals. Behavioral objectives. Contract learning.	Management Information Systems. NCHEMS Costing and Data Management System. HEPS (Higher Education Planning System.	Evaluation reports of "program side effects." "Holistic" evaluation of educational programs in the arts.
Basic Assumptions	Best evaluation is the expert opinion of a qualified professional. There is a commonly shared value system in the "arena" of the program or thing to be evaluated and/or. A "compromise mechanism" exists for accommodating differences in professional opinion.	Best evaluation is obtained from measurement data. Thing to be evaluated has measurable attributes. Instrument effectively measures attributes selected.	Best evaluation is based on an examination of achievement in light of goals or objectives. Goals or objectives exist and are identifiable. Attributes exist (and can be measured) that indicate relative success in achieving goals.	Best evaluation is one that serves decision makers in specific decision situations. Decision making processes are rational. Different types of decisions require different information. Many decisions are cyclic; systematic processes can be devised to support them.	Best evaluation highlights actual outcomes and/or concerns of constituents and sponsors. The "real" effects of a program or thing can be identified (as can the concerns of affected individuals.
"Advance Organizers"	Values may or may not be explicitly defined. Evaluator expected to be an information collector, synthesizer and judge.	"Norm-referenced," quantitative values. Formal setting required for application of the measurement instrument. Instrument must be validated, reliable, etc. Evaluator must be a measurement expert.	"Criterion-referenced" (goal oriented) values. Measurement technology commonly used within the context of performance vs. goal assessment. Evaluator may be expected to "judge" as well as measure.	"Decision-oriented" values. Information system methodology. Evaluator should be an information system specialist.	"Wider-context" values (selected by evaluator). "Holistic approach" (all contributive elements, etc. considered as they relate to each other). Evaluator must be skilled in human interaction and identification of concerns.
Nature of outcome/ Interpretation	Educated, personalized opinion (judgment of worth) of a qualified judge or panel of judges. Interpretation in form of a "pronouncement."	Number or set of numbers which can be compared to other numbers or a standardized scale, data, descriptive statistics. Interpretation in reference to norms.	Judgment of worth based on comparisons between performance data and objectives or standards of performance. Interpretation based on relative discrepancy or congruence.	"Continuous," timely, and relevant information for administrators to assist in judging decision alternatives. Interpretaion an administrative function; evaluator concerned with extraction and reduction of data.	Descriptive information regarding "actual" outcomes. Interpretation responsive to constituent concerns.

Figure 1: BASIC ASSUMPTIONS AND DISTINGUISHING CHARACTERISTICS. (Gardner (1977: 586-587).

	Professional Judgment	Measurement	Congruence Between Performance and Objectives	Decision Oriented	Goal-Free/ Responsive
Advantages	Easily implemented. Uses assimilative & integrative capabilities of human intellect. Recognizes outstanding expertise.	Proper validation and consistent application results in high comparability and replicability. Data mathematically manipulable. Generalizable results.	"Goal-orientation" provides objective basis for evaluation. Judgment criteria pre-established by objectives vs. performance measures selected. Relevant to current societal concerns.	Increased understanding of decision setting and information requirements. Focus on decision information needs assures relevancy of data. Encourages analysis of all factors affecting important decisions.	Flexible, adaptive approach. Useful in complex, relatively unstructured situations. All outcomes, etc. potentially relevant. "People-oriented" - high acceptance potential.
Disadvantages	Results criticized as non-replicable, non-comparable and overly subjective. Generalizability difficult or impossible.	Many variables difficult or impossible to measure. Often inappropriate and/or inflexible--serving available measurement tools instead of the "problem." Measurable attributes often irrelevant.	Focus may be too limited--not all worthy goals easily identified. Important side-effects may be overlooked. Tendency toward over-emphasis on end-product evaluation.	"Rational decision making" not predominant model in most real-life situations. In practice, frequent inability to cope with changing decision information needs. Inflexibility of "packaged" systems.	Relatively unstructured approach may be difficult to focus and manage. Results critized as non-replicable, non-comparable and overly subjective. Questionable credibility if evaluator non-expert in area of thing evaluated.

Figure 2: ADVANTAGES VS. DISADVANTAGES. (Gardner (1977: 586-587).

publication. The expert examines the thing (person, process) to be evaluated and then renders an expert opinion as to its quality, effectiveness, or efficiency. The statement of relative worth is the evaluation.

2. Evaluation as Measurement. This model assumes that the thing (person, process) to be measured has measurable attributes and that we have an adequate instrument. Examples that are familiar include SAT and other standardized tests and many attitude surveys. The evaluation manifests itself as a set of numbers which can be compared to some norm (other numbers, a standardized scale, or descriptive statistics).

3. Evaluation as the Congruence between Performance and Objectives. This approach assumes that there are goals and that we can determine the degree to which they have been achieved. Current examples of this model are the many competency-based methods of certification, contract grading, and systems based on behavioral objectives. The evaluation is some judgment of the relative discrepancy or congruence between stated goals and measured achievement of those goals, between objectives and performances.

4. Decision-Oriented Evaluation. This approach assumes that decisions follow from evaluations and that decision-making processes are rational. There are few good examples to point to in sociology, but industrial sociologists would be familiar with many management information systems being used today.

5. Goal-Free/Responsive Evaluation. "The central theme of this approach is critical examination of the institution, project, program, or thing irrespective of its goals" (Gardner, 1977: 583). This model assumes that the best evaluation highlights the actual outcomes and/or concerns of the constituents, as well as the sponsors, of the evaluation. It argues for an open-minded and sensitive approach to evaluation. The evaluation itself consists of a great deal of descriptive information, much like an ethnographic account, regarding actual outcomes, and provides an interpretation responsive to constituent concerns.

Having offered at least a beginning answer to these three questions, we can now move on to three larger questions. The first two of these pertain to the theme of this conference: knowledge needed, knowledge available. The third deals with the use of this knowledge.

My fourth question was: What do we know? In developing a tentative answer to this question, I will closely follow the work of Anthony Grasha (1977) who has provided the best single summary of the literature I have seen.

Grasha begins his discussion with a response to the question. Why evaluate? The answer consists of two major parts. First, it seems clear that evaluative response improves performance. To the extent that we are concerned about the performance of teachers and to the extent that evaluation results in feedback, teaching should improve. From Eble's work, Grasha finds other positive impacts of assessment programs (Grasha, 1977: 11):

> It can act as a catalyst to provoke campus-wide consideration of qualities that affect teaching and learning. Assessment might also attract into the profession those who see teaching as an attractive calling. Using students to help provide data encourages participation between teachers and students which, in large schools, may not be present in other forms. . . . /T/his involvement by student and teacher in the teaching-learning process can raise the level of instruction . . . Finally, an institution may be stimulated by the assessment of teaching to consider its overall educational goals, curriculum, and values.

Another side of the question, why evaluate? is: if evaluation is so positive, why haven't we done it before, why don't we do it more, or better? We have already alluded to the measurement problem. But a different kind of problem is that of faculty resistance. Evaluation of teacher-professionals is a sensitive, often emotional issue and frequently evokes defensive responses. Also, few people feel that their own teaching is in any way inadequate; the classroom is viewed as a private domain, and assessment is often seen as a potential threat to self-esteem and job security. To overcome this resistance, Grasha presents eight principles.

1. There is no perfect evaluation system. This does not mean that evaluation is impossible. It does suggest that "current procedures will improve faster if people are willing to assess the advantages and disadvantages of their practical experiences" (1977: 12).

2. Assessment activity should be broadly based. Teaching is a multidimensional activity. Too often, only one activity (classroom performance) and one data source (students) provide the basis for assessment. The entire range of teacher activities (including, e.g., teacher, advisor, researcher, professional contributions) should be included in assessment.

3. Evaluation must be linked to the reward structure of the institution. Simply put, the consequences of the assessment must be meaningful and be clearly understood by the person being assessed.

> An individual must know that assessment activity is supported and the degree to which teaching performance is used in assessing his value to a particular institutional unit. Otherwise, as examples from higher education show, people devote their energy to activities that traditionally count, e.g., committee work, research and publications.

4. Evaluation is best considered a process rather than a terminal event. The traditional way of gathering information is to survey students near or at the end of a course. This approach makes it difficult to review on-going activity. In our own research, none of us would use a single measurement when several are possible. The assessment of teacher effectiveness should not be a one-shot affair.

5. Evaluation systems should allow for the assessment of inter-individual differences and for the assessment of personal goals and objectives. Each teacher is unique, with unique goals. These should be taken into account.

6. Evaluation systems should maintain a personal and career growth and development focus.

> A system that only attempts to provide a data base for promotion, reappointment or tenure or to "hire and fire" is an empty system and eventually encounters resistance. Development must be an issue from the beginning to insure the long term survival of evaluation procedures (1977: 15).

7. People who are affected by a system should participate in its development.

8. Attempts to make personnel decisions on the basis of an objective quantification of assessment data should be avoided.

Grasha follows his discussion of overcoming faculty resistance with a summary of the "consistencies in the literature" on teacher evaluation. Although I do not present the references here, each of his conclusions is backed by published research. The literature contains both clear trends and contradictions. Among those which seems clear are the following.

1. Ratings of Teacher Characteristics are not Strongly Related to the Amount of Content Students Learn. (At least, where learning is measured by course grades or exam scores.) Some of this is due to the criterion problem. Most studies do use exam scores or grades as criteria. But a student's course of exam performance can be affected by many variables other than teaching, e.g., general academic aptitude, study skills, previous knowledge or experience, writing style, and grading biases.

Such rating systems, however, are not without merit. Within limits, student ratings are both valid and reliable. Student ratings seem to be no different than colleague ratings, in that the two sets of ratings are generally positively correlated statistically. Furthermore, there is a positive relationship between teacher ratings and both the likelihood of students meeting course objectives and taking further courses in the same area of study.

2. *Most Faculty and Student Demographic Characteristics Play a Minimal Role in Influencing the Ratings*.

> 1. There is no evidence in the literature that students attempt to damage an instructor through ratings. If anything, students are rather liberal in assessing performance. . . .
> 2. . . . /S_/ome evidence exists which suggests students *required* to take a course *may rate it somewhat lower than* those who elected to take it . . .
> 3. Class size usually has a small effect on *overall* performance rating . . .
> 4. There is virtually no relationship between ratings and sex of the instructor; sex of the student; academic grade level of the student; student grade point average or expected grade in a course; the research ability of the instructor; or the severity of grading.
> 5. Student ratings tend to improve with the experience of the faculty member. . . . /G_/raduate teaching assistants who had taken a course in teaching principles had ratings that were higher than other graduate teaching assistants. . . .
>
> (Grasha, 1977: 24-25)

3. *Teacher Characteristics which are Associated with Effective Teaching Have Been Identified*, at least from the student's point of view. Some of these are summarized in Table 1.

4. *Instructional Methods May Be More Effective than Instructor Characteristics in Influencing the Student's Learning*. Generally, methods which emphasize a high degree of student participation, collaborative problem solving and discussion skills, and opportunities for students to demonstrate independent thinking will be superior to the traditional lecture format. There are two major implications (1977: 29):

> One is that other things besides content achievement are potentially legitimate classroom outcomes. The other is that evaluation of content and non-content outcomes should focus on *the adequacy of the methods for meeting the obtained goals. Instructor accountability would rest on being able to demonstrate that the methods employed were able to produce the desired outcomes.*

BEST COPY AVAILABLE

Table 1

COMMON DIMENSIONS ASSOCIATED WITH TEACHER BEHAVIORS:
DIMENSION DESCRIPTIONS AND SAMPLE ITEMS[a]

A. Hildebrand and Wilson, 1970

1. Analytic/Synthetic Approach: The teacher knows the theoretical background and recent developments in the subject from several points of view. The ability to discuss a variety of ideas and to illustrate the relationship among them in order to increase the students' broad-based understanding of the area is stressed.

   Items: (a) Discusses points of view other than his own. (b) Contrasts implications of various theories. (c) Discusses recent developments in the field. (d) Presents origins of ideas and concepts.

2. Organization/Clarity: The teacher is diligent in planning the lectures with clear objectives to provide an organized picture of the information.

   Items: (a) Explains clearly. (b) Is well-prepared. (c) Gives lectures that are easy to outline. (d) States objectives for each class session.

3. Instructor/Group Interaction: The teacher allows students some responsibility in contributing to the educational process through discussion and the mutual sharing of ideas. The instructor is concerned about the class' response to his or her teaching.

   Items: (a) Encourages class discussion. (b) Invites students to share their knowledge and experiences. (c) Invites criticism of his own ideas. (d) Knows if the class is understanding him or not.

4. Instructor/Individual Student Interaction: The teacher is approachable, interested in students and treats them with respect.

   Items: (a) Has a genuine interest in students. (b) Relates to students as individuals. (c) Is accessible to students out of class. (d) Respects students as persons.

5. Dynamism/Enthusiasm: The instructor is energetic and stimulating, conveying the impression that he or she enjoys "turning students on" to his subject.

   Items: (a) Is a dynamic and energetic person. (b) Seems to enjoy teaching. (c) Is enthusiastic about his subject. (d) Seems to have self-confidence.

B. Isaacson, McKeachie, and Milholland, 1964

1. General Teaching Ability: Teaching abilities that tend to form a consistent pattern across different instructional styles.

   Items: (a) Material is put across in an interesting way. (b) The intellectual curiosity of the students is stimulated. (c) Things are clearly explained. (d) The teacher is skillful in observing student reactions.

2. Overload: The amount and difficulty of the work assigned.

   Items: (a) Assigned very difficult reading. (b) Asked for more than students could get done. (c) Assigned a great deal of reading.

3. Structure: The ability of the instructor to organize and plan a course.

   Items: (a) The instructor decided in detail what should be done and how it should be done. (b) Followed an outline closely. (c) Had everything going according to schedule. (d) Planned the activities of each class period in detail.

4. Quality: The concern the teacher has for the quality of classroom related work.

   Items: (a) Told students when they had done a particularly good job. (b) Complimented students in front of others. (c) Criticized poor work.

5. Group Interaction: The nature of the interpersonal interactions in the classroom.

   Items: (a) The students in this class were friendly. (b) In class, I felt free to express my opinions. (c) Students argue with one another or with the instructor, not necessarily with hostility. (d) The students frequently volunteered their own opinions.

6. Student-Teacher Rapport: The nature and quality of the interaction between student and teacher.

   Items: (a) The teacher listened attentively to what class members had to say. (b) The teacher was friendly. (c) The teacher was permissive and flexible. (d) The teacher explained the reasons for his criticism.

---

[a] Source: Grasha (1977: 26-27).

My fifth question was: what do we yet need to know? A tentative answer to this question is presented elsewhere in this volume, in the chapter by Wagenaar.

Finally, we can ask how we shall use our knowledge to change teachers. This fundamental question raises others: How do teachers become better? What are the consequences of change? What are the institutional, professional, and disciplinary constraints? Many of these kinds of questions are being addressed. For example, the ASA's Projects on Teaching, through its surveys, workshops, publications, and services, has done much to identify needed change to implement such changes in individuals, departments, and even the Association itself.

References

GARDNER, D. (1977) "Five evaluation frameworks." Journal of Higher Education 48: 573.

GRASHA, A. (1977) Assessing and Developing Faculty Performance: Principles and Models. Cincinnati: Communication and Education Associates.

HILDEBRAND, M. and R. C. WILSON (1970) "Effective university teaching and its evaluation," in K. Eble (ed.) The Recognition and Evaluation of Teaching. Salt Lake City: Project to Improve College Teaching.

ISAACSON, R., MCKEACHIE, W., MILHOLLAND, J., LIN, Y., HOFELLER, M., BAERWALDT, J., and K. ZINN (1964) "Dimensions of student evaluations of teaching." Journal of Educational Psychology 55: 344-351.

Assessing Teacher Effectiveness in Higher Education: Why, When, Where, and Perhaps Even How?

Gerald Marwell

University of Wisconsin - Madison

Our assigned task is to consider what we know and what we need to know about the assessment of teacher effectiveness, particularly as it pertains to the teaching of sociology in universities and colleges. I am specifically charged with consideration of the literature published prior to 1972. On consideration, I take that charge as a license to attend mostly to the problem of what we do not know. Any knowledge developed prior to 1972 should certainly be reflected in the later literature. In what is supposed to be a cumulative discipline, the early good should have been recognized and built upon, the early mistakes recognized and corrected. My literature review led me to no hidden jewels of ancient wisdom, although it did lead me to one article described by Hansen and Kelley (1973) as "the best summary of research on course evaluations" (Costin, Greenough, and Menges, 1971). Without seeing what has been done recently, I would agree with that assessment. It is a very good effort of its kind.

Like most articles in this area, though, the one by Costin, Greenough, and Menges is directed primarily towards how to evaluate teaching, and is based on the notion that we are engaged in a fairly typical measurement task. Given the quality of the research reviewed, and the high-powered talent that has been devoted to the task of measuring teacher performance since this article appeared, I am persuaded that we know well enough how to assess teacher effectiveness--given certain assumptions. We could probably choose among several strategies which would yield quite reliable and even valid assessments. What we may not know very well is why we are performing this expensive and sometimes odious task, and under what conditions it should or should not be undertaken.

Teacher Evaluation and the Improvement of Teaching

Presumably, the assessment of teaching is not a goal in itself; it is a means towards an end. While other goals might enter consideration, the one most of us would argue for would simply be the improvement of teaching. Yet, as far as I can tell, there is no evidence available to show that the assessment of teachers leads to the improvement of teaching. Although the relationship seems self-evident to many, I believe a good case can be made for the proposition that teacher evaluation in most colleges and universities leads to a decline in teacher quality. At the very least, I believe the null hypothesis of no difference is far from disproven.

We can probably summarize the processes by which evaluation is supposed to improve teaching in four major points:

1. Selection: If we fire the bad teachers, and encourage the good ones not to quit, we will surely raise the average. Good evaluation would allow us to know which is which and apply the policy correctly.

2. Motivation: If we reward people more when they teach well they will be motivated to work at their teaching, and thus generally become better. Good evaluation correctly specifies who should get whet when.

3. Norms: The very presence of evaluation, some argue, establishes departmental and college norms which emphasize that good teaching is valued. This, in turn, encourages conformity to communal expectations.

4. Feedback: Some teachers do not know how their behavior is affecting students. If they knew, they would change their behavior to get better effects. Evaluation, when done correctly, provides such feedback, hence producing improvement.

It all sounds fairly reasonable. But is it accurate? How do things proceed in the contaminated reality of contemporary academe?

1. Selection: There is probably an enormous amount of variation among institutions in the effectiveness of evaluation in weeding out poor teachers.

In many institutions, publication and scholarly work are such an overriding concern that teaching evaluations have to be exceptionally poor to make any difference. Brilliant teaching evaluations might keep someone a job if his scholarly work is marginal. Since the vast majority of us, however, are near average (I still believe in the normal distribution of such talents), selection is probably not very much influenced by teaching evaluations. The old saying, publish or perish, is not about teaching competence (at least directly).

In institutions which are less oriented toward publication as the measure of performance, teacher evaluations might be very important for selection. However, a number of factors might vitiate these effects even here. Many such institutions are unionized, and unions show little sympathy for evaluation of performance as an important determinant of retention.

Faculties generally protect their own. They are likely to use student evaluations of teaching as little as possible in making their decisions. Consider as relevant evidence the role of teacher evaluations in high schools. Here, it seems, assessment has, at best, marginal impact on the retention process. Why should it be much different in institutions of higher education?

Before leaving the question of selection, two additional points are worth noting. The first is that selection on teaching competence probably works best if information is available at the time of hiring. Thus, the evaluation of young teachers might affect whether (and where) they are hired if they are not retained by their present school. Similarly, evaluation may be effective if it is applied to graduate students who teach at their training institution prior to going on the job market. Although the institutions at which the evaluations are done may not benefit by these evaluations, the system as a whole may very well find them useful for selection.

Within institutions the assessment of teaching might also be useful for the assignment of teachers to classes. Some teachers might have talents for certain types of classes or topics but not others. Or poor teachers might simply be placed where they can do the least harm, while good teachers are rewarded (ha) with large classes where they can do the most good. Whether this process works in the real world to have a positive effect on teaching depends in large part on the degrees of freedom available to an institution by virtue of the size of its staff and the diversity of its offerings.

2. Motivation: If anything, the evidence here is likely to be even more difficult to gather. I suspect that the relationship between assessment and motivation might vary over time as well as over institutions. My personal experience suggests that when a system of teaching evaluation is introduced in a given department the effort put into teaching does increase. However, this increment of effort declines over the next few years, perhaps eventually disappearing altogether. The decline in effort may be traced to many factors. Just the realization that people don't take the evaluations very seriously might be important. More tragic, however, might be the fact that many teachers do not find added effort resulting in better assessments. We are still taking as an article of faith the notion that there is some direct return to effort invested in teaching. Such faith has not been demonstrated to be well placed. The tragedy here may be that teachers who try

and fail become hostile to students, alienated from their work, and even worse teachers than previously. At the other extreme we have teachers who receive good ratings. Knowing they are secure in this area they might reduce the effort put into this part of their job. Although their lectures remain popular they might not work at keeping them fresh. Have the ratings helped their motivation?

3. Norms: Sociologists have long noted the difficulty of changing norms through changes in formal rules. Prohibition didn't stop people from drinking, for example. Why should the institution of evaluation be any different? Do the members of the faculty believe in the legitimacy of teaching assessment and the particular process of assessment chosen? This might well be the deciding factor in knowing whether assessment will be able to change norms. But we have lots of reason to think that many, perhaps most faculty will not approve of evaluation. Instead, they might consider it an illegitimate infringement on their freedom in classes, and a threat to their security and well-being as well. Even where assessment is generally approved in principle it is difficult to come up with a concrete system accepted by all--or perhaps even by most--as accurate and useful.

4. Feedback: In some ways this is the crucial factor in the relationship between assessment and teaching quality. It is the factor upon which the most pious among us rely. We are not interested in punishing poor teaching, but we still believe that most teachers want to be both good at their trade and well-liked by their students. Thus, if we only inform them where they are being badly received . . . The notion here is that information, by itself, without rewards and punishments directly attached, will affect behavior. But will it? Again, it does seem to me we have reasonable information. My personal observations are that feedback is frequently useful to new teachers. They both understand that they lack experience and have a useful excuse for any problems that become evident with their teaching. They tend to use feedback less defensively than older professors. At least for one semester. Beyond that, though, the utility of feedback is questionable. By the third or fourth time you've heard the same complaint, you know, damn it, you know! You also know how difficult, or even impossible it would be to change--at least without major damage to other enterprises, or aspects of our lives.

A rather relevant psychological principle to consider in the matter of feedback is that nobody likes the bearer of bad news. In the case of teacher evaluations this sets up some very unhappy situations. If one relies on student evaluations, professors who receive less enthusiastic support than they expected (which might even be average or good ratings for professors who think they are superb teachers) might become angry with the students, who essentially bring them the bad news. Or angry with the process of evaluation. Neither reaction is conducive to an improvement in teaching and might make the professor punitive, or otherwise less desirable as a teacher. Similarly, ratings by colleagues, even when temperate, generous, and oriented towards helping may often be clear disconfirmations of the teacher's self-vision. Such negative messages might mean acrimony and distrust among colleagues, feelings of guilt or even resentment on the part of the evaluator. None of these is an outcome to be sought within departments.

Some Additional Costs of Evaluation

The reason that we need to have some idea that assessment actually improves teaching is that the assessment process can be so damn costly. I hate to be practical--it so rarely befits an academic--but the money, time, and energy spent on evaluation are hardly trivial in an era when funds are harder and harder to come by, time has become an enemy, and the energy crisis is on every front page. My personal impression is that we currently put much too much time, effort, and resources into teacher evaluation for the payoffs we receive. I have argued above that current practices of evaluation may be contributing little. Moreover, I gravely doubt that further investment in the design of teacher evaluation can substantially improve the situation. Yet there are groups all over the country expending the time and energy of highly paid professionals, in committees, to do just that.

Let's face it. Teacher evaluation is very expensive. In my department we use student questionnaires as our primary method. For paper, computer time, payment to key-punchers, and one person who spends a considerable amount of time organizing the materials, the cost to our department probably runs close to $4000.00 per semester. Half of a professor's semester salary. We might have been able to do a better job of improving teaching by hiring another half of a professor and reducing class size or offering a greater variety of courses among which students could choose. Where classroom visits are used for evaluation, the costs are of a different nature, but just as real. I don't know about the rest of you, but to me time is worth lots of money, and I would not be happy about spending my time in other professors' classes for the purposes of judging how well they were doing their job. Especially if my judgment is negative--because I know I am in for some difficult hours. If my judgment is positive I suspect I've wasted my time anyway.

The more we orient our evaluation programs toward improving teaching, the more expensive it gets. If we use student evaluations we must begin to ask a large number of questions aimed at finding specific problems with the teacher's methods. The more questions one asks students, the more costly the procedure. If we use observation, we must spend more and more time with the teachers and then spend additional time trying to show them the right way to do things.

The Politics of Teacher Assessment

With all these negative attitudes among faculty toward the assessment of teaching, and with the practical costs of assessment far from insignificant, why is assessment so widespread and growing? Perhaps the popularity of the practice argues empirically against the truth of the notions advanced above? Might it not be that despite the possible shortcomings of the system there are advantages enough for people to want the system?

This, I feel, turns out to be a sociological question rather than one of practical effectiveness. When I look at the college and university system of the 1950s--growing, respected, understaffed, valued--I see a system with a relatively small amount of teacher evaluation. In institutions where the faculty are recruited as scholars, where they are secure, where their value is not questioned, the arguments I have presented are strong enough to keep assessment a matter of individual choice or specific situation. The campus revolution of the 1960s was accompanied by increased outside interference in the affairs of institutions of higher education, a public devaluing of the enterprise, the demystification and reduced emoluments of the professoriate, an enormously increased supply of certified teaching labor, etc. These all brought with them increases in the amount of teaching assessment. The relationship does not seem accidental. Teacher evaluation, I think, is more a response to political pressures, than a move towards professionalization and rationalization. The cry has been for accountability--and it has come largely, if not exclusively, from outside the faculty itself.

Not that there are no interests within the faculty that are served by the evaluation of teaching. Again, we do not have data regarding attitudes towards teacher evaluation, but a fairly crude interest theory might suggest that attitudes would depend on both the individual's level of research productivity and expected level of teaching assessment, as described in the following:

TEACHER RATINGS	RESEARCH	
	Non-productive	Productive
Good	Yes	Ambivalent or No
Medium	Yes	No
Poor	No	No

My version of interest theory makes fairly obvious predictions for most cells. The exception is in the "productive-good ratings" cell, where I see ambivalence or negative attitudes. Here, I suspect that these very secure professors would see teacher evaluation as a threat to their ability to keep colleagues whom they value for their intellectual and research capacities, but who might be poor teachers. In this conflict of values I guess there would be a general trend toward rewarding research more than teaching.

Where there are differences in interests we might expect conflict, with the group with the most power winning. In research-oriented institutions we might expect most of the highly productive persons of the past to have survived, regardless of teaching competence, while less productive good teachers may or may not have perished. The balance of power should be against evaluation. Even in more teaching-oriented departments the research-productive tend to survive, and they may do battle with the equally surviving good teachers. Only the unproductive poor teachers perish, but this makes the battle much more equal. When outside political forces or emergent student power impinge on this situation, the balance is decisively shifted toward assessment.

The groups that want assessment are groups that want to have an impact on the teaching process, and feel they do not have leverage in other ways. Students want to punish, or inform, or even help teachers they don't learn from, or enjoy, or like. Administrators want to increase satisfaction from clients--be they students, parents, legislators, or contributors--and also improve teaching. They can use evaluations as reasons for decisions oriented toward teaching as a value. Legislators, parents, et al. share the administrative viewpoint and lately are even more likely to see professors as service-providing employees rather than as scholars. Assessments mean the criteria and justifications for action in those terms, rather than the scholarly ones held by much of the faculty.

Political Evaluation

The picture with which I am left is one of conflict among the faculty--us--and everyone else. If that is the case, perhaps we should consider teacher assessment in this political light instead of solely in terms of its educational utility. How, we might ask, can evaluation be performed as a useful political act, one which protects us from our critics. How can we design our procedures to maximize their political utility and, incidentally, reduce their cost?

Perhaps the best strategy to employ would be one that makes the various interest groups feel they are having an impact on the system. It is this process that particularly recommends student ratings as a means of teacher assessment. Observation by other faculty has a distinctively in-group and potentially corrupt image. The criminals are also the police. One faculty member scratches the other's back, and expects the same in return. On the other hand, the other interest groups (regents, parents, et al.) know that students are the basic clients for teachers. They are there, observing the teacher, as part of their basic routine. Lord knows parents, regents, trustees, and newspapermen have no interest in themselves spending time listening to some professor talk, certainly not to someone who isn't a current media hero, at any rate. So the students are convenient representatives of the coalition of hostile interest groups.

Besides, everyone knows that students are a surly, hostile, uncivil, anti-establishment, rebellious, tactless, negative and generally frightening lot. Not the kind of people one would pick for most jobs, but getting (I beg your pardon, evaluating) the faculty might be just up their alley. If they won't stick it to the faculty, nobody will.

Interestingly enough, students share this vision of themselves and their peers. They also believe themselves tough, and the best people to strip the facade of competence from the charlatans in the classroom. Even those with more moderate visions fervently believe that as the teachee they are the logical and obvious person to accurately gauge the competence of the teacher. So all (but you and me) seem to agree on the students as the evaluators.

In the face of unanimity it is politic to accept gracefully. Because students may really be best from our point of view as well. For the average student behaves little like his public image. Sure, there will be some firebrands in any class who may want to say things just to fit hostile assumptions. But we never look at individual student evaluations--just averages over whole classes. And the hostility of the few is overwhelmed by the general geniality, or apathy, or confusion of the many. We are, after all, authority figures to most students. Their basic response (one which we often detest and work hard to overcome) is more awe than awful. Getting them to disagree with us takes talent and application on our part. They believe our credentials--which is why they show up in the first place.

If you look at the ratings which typically emerge from students you find that the average teacher is rated as "better than average." This seeming impossibility has sound psychological roots. First, a reasonable number of students stay away from teachers and courses they don't want to take, and select teachers they want. Thus, you do not have a wholly random set of students in your course, and they will tend to be more positive than would a random group. This is emphasized by the fact that students who don't like you will often have dropped by the time of evaluation or are most likely to cut that class.

More interestingly, concrete individual teachers are human beings with redeeming qualities to the student, while the "average" teacher to whom he is being compared is a conglomerate of memories and stories from other students. These stories tend to emphasize complaints and weird behavior, not tales of fabulous performance. Unfortunately, our most salient meaning to most students lies in the power we hold over their futures and the threatening possibility that we may fail or otherwise hurt them. When we do they are likely to seek solace and help from their friends and spread the news, often justifying themselves by embellishing their side of the case. Thus, the collectivity looks worse to a student than any reasonable average of real teachers.

Tactics of Evaluation

There are a few tactical lessons in the above paragraphs. First, never give your evaluations too early in the semester. Wait for malcontents to drop. On the other hand, don't give them too late. While hope for the good grade still springs your ratings will be better. Never warn students beforehand that evaluations will be given. Your detractors might show up just to get you. On a normal day the ones missing from class will be heavily weighted with the unhappy. Always point out that these evaluations will have real effects on your salary and career. Students, like most other people, tend to be kind, and don't want the moral responsibility of hurting you. Always compare yourself to some mythical average professor at the University, not to some concrete other person or absolute standard.

Another good strategy is to encourage the students to write long essays on you. This makes them feel they can really get it off their chest and "communicate." But essays can never be summarized in precise enough fashion to be used against you. On the other hand, no matter how bad you are there are always some positive statements and these can be culled to support the notion that at least some students love you when you come up before a committee reviewing your status.

Side Benefits

And there is still more. First, a genuinely political view of teacher evaluation allows us--even requires us--to spend very little time and effort designing our evaluation instruments so that they are particularly valid or reliable. We want there to be a considerable amount of fuzziness in the statements we can make from evaluation. For example, it is not in any department's interest to be able to say with certainty that Professor A is just below average as a teacher, while Professor B is just above average. These invidious distinctions are not conducive to collegial relations within the department. Much better is a system in which everyone within a substantial number of ranks from one another can feel that the difference is just random error. That way the teacher ratings legitimately enter such decisions as salary differences and promotion only for those who are substantially different from everyone else. They must be very, very unpopular, or very, very popular.

Since the major way in which psychologists increase the reliability of their measuring instruments is to increase the number of items which they then combine into an index, poor-quality evaluation means short questionnaires.

This allows us to save money. Never use more than one question where only one will do. The best evaluation thus might be restricted to two questions: "Compared to the average Professor at this university, how would you rate the teacher in this course? 1. Excellent, 2. Good, 3. Average, 4. Fair, 5. Poor"; and "What grade do you expect to get in this course?" This would be followed by a space for a long essay on exactly what is good or bad about the teacher and suggestions for improvement.

From where, you may ask, does the second question come? I have already explained the first question and the essay, but why ask about expected grades? Part of the reason is obvious. We all know that easy graders are more popular with students than hard graders. Again, the correlation is far from perfect, but the non-threatening professor who is an easy B will have far fewer strong negatives than the tough nut. All of our political audiences understand this principle.

And the principle provides us with an interesting possibility. Suppose you are being roundly castigated by your students as a poor teacher. One response you might consider would be to raise your grades, thereby improving your ratings. But you would undoubtedly be making a political error. Your colleagues and others will note the ploy and dislike you. It might not even work. And if everyone competed in this way, administrators, trustees, et al. would become exercised.

Consider the alternative. Become a tough grader. With this system, you provide a legitimate explanation for your poor ratings. You will be able to show that students who get good grades like you, and claim that the reasons so many students dislike you is that you are not one of these easy graders who seems to abound in our colleges and universities today. You will not buy popularity at the expense of our time-honored standards. Everyone can understand that. You might also encourage the more popular teachers to improve their grades, pointing out that they get the best students, and that students who learn more should get better grades.

On Second Thought

Reading the above, I wonder how seriously I mean it to be taken. Although the logic of the argument seems to roll an over-more-cynical snowball, the final position may be too radical even for me. I am particularly concerned that some may interpret the above as saying that teacher evaluation is hypocrisy, that we are really politically motivated, and just pretend to be trying to improve teaching. Therefore, we should stop being hypocrites and stop evaluating.

This isn't really my position. That I have profound ambivalences about the usefulness of teacher evaluation should be obvious. I think we should fully understand the potential costs of the process to ourselves and our profession, as well as aspire to the potential benefits. And part of the understanding we should have is that there are myths abroad which are inimical to the continued well-being of the organizations in which we have so much invested. The political use of teacher evaluation is not hypocritical per se. It is a matter of promoting external understanding through communication. And of lining our pockets with some of the proverbial silver lining as well.

The Bureaucratic and Disciplinary Contexts of Teaching

Gail Brady
University of Missouri
Ted K. Bradshaw
University of California - Berkeley
Zelda Gamson
University of Michigan
Hans O. Mauksch
University of Missouri
Reece McGee
Purdue University
Lawrence J. Rhoades
American Sociological Association

Social changes in the wider society are exerting pressures on higher education to become more accountable for the behavior of teachers and the learning of students. These pressures are being exerted by various levels of government, by professional associations, employers, students, and by the general public in response to demographic changes in the traditional student population brought on by lower birth rates, the equalization of educational opportunity, and labor market conditions.

The traditional population from which higher education drew its students until recently were young people between the ages of eighteen and twenty-two. The size of this pool has declined over the past decade because of lower birth rates, the elimination of the draft, and the declining belief in the relationship between higher education and economic success.

This wane in the traditional population of higher education has been offset somewhat by pressures for the equalization of educational opportunities. The push for equality has changed the composition of the student population in a number of ways. First, it has led to increased college attendance among women and minority groups. Second, it has expanded age boundaries, as more older people are attending colleges and universities. Overall, egalitarian pressures and practices have introduced greater heterogeneity into the student population in values, ethnicity, academic skills, and social class.

Labor market conditions effect higher education because they shape the careers of students and faculty members. This impact is caused primarily by the inability of the economic structure to produce the number of positions needed to accommodate large numbers of college graduates. Because of competition for jobs, students demand courses and instruction that will equip them with the skills they need for employment. Since teachers also face a tight labor market that severely restricts their mobility, they become vulnerable to pressures from students and institutions. These pressures are likely to have a powerful impact on teaching. In order to understand the effects of these forces on the teaching of sociology, we must look at the way teaching is shaped by the organization in which it is performed and by the norms and authority of the discipline which determine its contents.

Bureaucratic and Disciplinary Systems

A key consideration in any assessment of the academic profession is that, unlike the traditional free professions such as law or medicine, academic people work within a framework of double allegiances and norms, both of which profoundly influence them and the conditions of their activity. The two halves of the framework are the educational organization which employs them, and the discipline to which they belong. At various points in the individual's life and work, the norms, goals, and activities of these two will be in harmony; in others they will be incongruent or even in conflict.

The employing organization--the college or university--is an observable social entity. It has a geographic location, a table of organization, employees, buildings, rules, an authority apparatus, and so forth. The discipline, on the other hand, is more an ideal than an organization. It is manifested in the professional association, with its officials, constitution and by-laws, meetings, reports, and periodicals. It has a membership roster, code of ethics, traditions, conventions, norms and values, a stratification order of elites, journeymen, and an obscure definition of success and failure.

The discipline is embodied within the college or university primarily in the department which, in the American system, exists to transmit through teaching the accumulated knowledge associated with the discipline and to further that knowledge through research. The department, thus, is a key focus for sociological analysis of the academic profession because it intersects the bureaucratic and disciplinary systems. It often functions as a marketplace where values from one system are traded off or bargained against resources from the other. It is specifically within this marketplace function that the influence of the two may be seen in highly specific teaching problems.

We approach this formulation through the concept of authority structures. The authority structure of the organization is bureaucratic; the authority structure of the discipline is collegial. The authority of the bureaucracy is legitimized by the state and, in the case of some private institutions, by some religious entity as well. The authority of the discipline is legitimized by tradition, academic consensus, and to a minor degree by the state which recognizes its incorporation. The organization needs the department in order to legitimize its activities; any college or university must offer courses in most of the traditionally recognized disciplines. The discipline, as represented by the department, needs the organization in order to carry on its activities as well; the formal transmission of subject matter must be conducted within the framework of a school. As a consequence of this interdependence, the organization provides resources the discipline requires in order to carry on its work, while the discipline provides the work activities that are the reason for the organization's existence. Thus to a considerable degree, the goals and activities of the two systems are congruent.

But to some degree, they are not. The principal goal of any academic discipline is the creation of knowledge; the principal goal of any American school, at least, is the transmission of knowledge. Further, institutional resources are finite; disciplinary needs are, at least in principle, infinite. Disciplines have reward systems for their members which do not necessarily coincide with activities required for institutional maintenance; institutional requirements do not necessarily produce the kinds of activities that lead to disciplinary rewards. The resources, rewards and sanctions, and authority structures of the two are to a significant degree non-overlapping. All of these issues intersect in the department where they express themselves as tensions not only in the individual careers of faculty members, but in routine departmental decisions such as setting student requirements, curriculum, hiring and promotion, resource allocation, and staff assistance. This is the sociological framework within which the context of teaching must be understood.

The Generation and Allocation of Resources

Most of the hard resources for teaching are within the control of the organization through external and internal funds. These cover faculty salaries, equipment, space, and staff assistance (teaching assistants, secretarial, administrative). The organization also controls the resources for institutional program development and planning, and it provides resources for support of teaching and inter-departmental/administrative coordination.

Disciplines, on the other hand, do not control hard resources. The principal resource available to the discipline is status, which becomes a means of access to internal and external resources. Status is conferred and confirmed by peer review--mostly from outside the organization--certification, publication in valued journals, honors of various kinds, and formal roles within disciplinary associations. Status from these sources can be exchanged for internal resources--a position, released time, higher salaries--some of which can be used toward enhancing teaching. Often, however, status-based resources are conferred to support research and scholarly activities. Beyond conferring status, the disciplines can provide some support and coordination for teaching (witness the projects on teaching that have burgeoned in disciplinary associations), as well as resources for program development. Until now, however, these have been limited to information and consultation.

The department has limited control over the amount of hard resources at its disposal, although it has some discretion about the way it will distribute those resources it receives from the organization--especially in the content, structure, and assignment of teaching. Some departments do generate their own resources from external funding for research; occasionally, there are funds for teaching improvement from external funding agencies. While the department generates only small amounts of money over which it has control, it does produce outputs that can be exchanged within the organization for money or goods and resources which money can buy (i.e., positions, space, equipment). Enrollments have increasingly become a medium of exchange within colleges and universities. Sometimes, an entire curriculum is seen as valuable for students and for the organization as a whole. In some colleges and universities, faculty output in the form of student performance, and in others, research publication are also used as counters for capturing hard resources.

The resources under the control of the organization and these few generated by departments have led to an elaborate structure of rewards and sanctions operating across and between both levels. We have come to see this issue as being at the heart of the bureaucratic and disciplinary context as resources affect teaching; they are critical to the motivation and performance of faculty in their teaching roles.

Bureaucratic and Disciplinary Rewards and Sanctions

It is useful to categorize the kinds of rewards and sanctions available to the organization as either formal or informal. Formal systems are to be found, and are probably used in a similar manner, at all types of academic organizations: universities, four-year colleges, and community colleges. The degree to which informal systems are applied by the institution may vary, with informal mechanisms employed more frequently in schools of lower academic prestige.

The two most prominent formal mechanisms are promotion and pay. Less visible, and less commonly used devices include named chairs, honorific titles, awards, and other symbols. Informal mechanisms commonly consist of manipulations of bureaucratic authority and regulations: award or denial of leave, allocation of travel money, research assistance, secretarial help, office equipment, and the hundreds of other matters normally left to the discretion and authority of presidents, deans, and department chairpersons. Sometimes what are here called informal devices may be inserted into the formal system, as when an institutional officer intervenes in the promotion process or salary proposals in ways permitted by the system, but not normally used.

Formal organizational sanctions are probably typified by the withholding of promotion or salary increases or, in the case of the non-tenured, simply by non-reappointment. Today, the latter may in effect constitute discharge from the academic profession in many fields; so it becomes a very great threat indeed. Formal sanctions available to any organization for the tenured also include oral or written rebuke from administrative officers and, in rare instances, discharge for cause as recognized by the AAUP. Instances of the latter occurrence are almost unknown, and discharge is probably more usually effected informally, through the exertion of organizational pressure for resignation.

One variation in the system, which may to some degree distinguish schools is the kinds of activity that are rewarded or sanctioned. Large universities, for example, traditionally ignore the private lives of faculty members; smaller schools may or may not consider off-campus personal behavior relevant for reward or sanction. A second related and more subtle phenomenon regards promotion and salaries. While formal rewards such as promotion and pay are within the organizational sphere, schools of differing types differ in the degree to which they apply them according to disciplinary or organizational criteria of rewardability. Thus a major university may feel compelled to reward an individual for research accomplishments that are lauded by his or her discipline for fear of losing the person to another institution, whereas the same scholarly behavior conducted in a community college might earn no recognition whatsoever, or even be sanctioned by the school administration as not being what the person was hired to do. Similarly, it is not uncommon for universities to be embarrassed when junior faculty recipients of excellence in teaching awards are dismissed by promotion committees because of low scholarly productivity.

As these examples suggest, while the actual mechanisms for employing rewards and sanctions through promotion and pay are similar in a variety of organizations, the criteria for their use vary by type of school. Finally, it should be noted that while salary ranges by rank are similar in most organizations regardless of type, there are often considerable degrees of overlap across ranks. This results from the fact that the two, while correlated, are not structurally related as they are in industry or the military. Academic rank is essentially honorific and is not tied to work assignment. In industrial and military organizations, a change in rank is usually accompanied by a change in duties or work assignment. This is not the case in the academic profession. The senior professor facing retirement has duties identical to those held during the first year as a college instructor.

A second variation in the formal system of rewards and sanctions is the level at which they operate. In general, this, too, is probably related to type of school, with the more prestigious (those typically more imbued with disciplinary values), tending toward a greater degree of collegial participation and authority. In major universities, departmental recommendations on individual salaries (whether initiated by the chair or a faculty committee) may for all practical purposes be the last word because the chairperson has already negotiated the total budget package. In small colleges, on the other hand, it is not uncommon for the Dean of Faculty or even the President to make salary decisions without any significant faculty consultation. Promotion decisions seem to run between the same poles, although some collegial recommendation, at least, is probably more clearly the norm, a fact perhaps not unrelated to the honorific meaning of promotion.

The most fundamental conflict between organizational and disciplinary formal rewards and sanctions occurs most commonly at the extremes of the prestige hierarchy. At major universities, the dominant criteria for promotion and salary increases are disciplinary. It is not uncommon for individuals to be highly rewarded by their organizations for contributions to their disciplines even though they may have ignored organizational responsibilities such as undergraduate instruction and committee work. At the other pole, in the community colleges or less prestigious four-year colleges, reward is more likely to be given for service to the school and for teaching activities; scholarly activity may even be seen as a distraction.

These conflicts derive from the larger social systems within which colleges and universities operate. The major universities, by historical accident, are committed to the pursuit and advancement of knowledge through research. This means they must compete with one another for eminent scholars in order to maintain their prestige. Many state-supported universities of this type also have legislative mandates for the instruction of undergraduates and for public service, but these are generally seen as burdens. It must reward success in the disciplinary-defined task of research in order to be successful; were a university to demand that its faculty pay primary attention to organizational needs, it would find itself unable to attract eminent faculty members.

By contrast, smaller schools of low prestige usually have civil or religious mandates for the instruction of undergraduates. Thus, they must demand that their faculty members focus primary attention on organizational requirements. This means that disciplinary values often go unheeded and disciplinary rewards, even if attainable, become counterfeit coin on campus.

Disciplinary Rewards and Sanctions

The discipline does not directly bestow rewards or impose sanctions, although the professional association which embodies it may do so in minor ways. The fact that a discipline has no structural existence, however, creates an interesting situation: in effect it depends on the college or university to bestow its principal rewards. But the principal reward the discipline has for individuals is status--based on professional, not organizational, values. People derive professional status from others' opinions of them, and they are rewarded by their organization in the form of promotion, pay, and emoluments. The high opinion a chairperson has of an individual's professional competence is translated into promotion and salary recommendations which are acted upon by organizational officers.

Again, it is useful to conceive rewards and sanctions as dichotomized into formal and informal categories. The disciplinary or professional association also has a formal reward and sanction system. Rewards from this source are largely honorific, consisting of election or appointment to various positions within the association, prizes for outstanding work, and temporary fellowships. Most disciplinary associations also have constitutional provisions for formal sanction after organizational inquiry, but the application of these is extremely rare. But the discipline, operating indirectly through the values and decisions of its members, has at its disposal powerful informal rewards and sanctions. Individual members serve as referees for its journals, advisors to the federal government, judges of research proposals submitted for funding, members of institutional hiring committees, authors of letters of recommendation, consultants to publishers, and evaluators of faculty at other institutions. In dozens of other capacities, individuals have the opportunity to make judgments of and decisions about others. Thus someone whose professional competence is low is likely to find it more difficult to publish than do others with greater talent and skill and, in consequence, in securing promotion or salary increases, placement on professional association programs, and funding for research.

The fact that the system is informal permits abuse, and it seems probable that most instances of abuse are never known to anyone but their initiator, since the motives for decisions are usually not made public. (Indeed, in many instances, even the identity of the decision-maker remains unknown to the subject of the decision.) Again, as with the formal system of disciplinary reward and sanction, it is often other organizations which act as agents for the discipline: to hire or not to hire, to grant or to withhold funding, to publish or to reject an article or a book.

Congruence and Incongruence Between the Organization and the Discipline

We noted earlier that there is some reciprocity between discipline and organization embodied in the academic department. The discipline exists for the creation of knowledge, the college for its dissemination. Each must have the other in order to operate. But the mutuality of interests between the two is far from perfect. On the one hand, disciplinary prestige is based on activities that often have little direct relation to immediate organizational functions. On the other hand, many of the activities essential to the functioning of the organization contribute little or nothing to the advancement of the discipline or to a person's status within it. At least in the major research universities, there is a difference between what faculty members are paid to do by the organization and the status they receive from their disciplines.

These circumstances create the context in which conflicts arise. Organizational rewards are finite; certain activities must be undertaken and satisfactorily performed if the institution is to function--teaching and some committee work, for example. To some degree, then, these activities must be rewarded. But for academics involved in running their organizations, values may be more important and the activities supporting them more worthy of reward. The contending organizational and disciplinary activities require efforts that are to some extent contradictory. The professor gathering material in the library cannot simultaneously be in the classroom, and hours spent on the student admissions committee cannot simultaneously be spent in Washington evaluating funding proposals.

These tensions and contradictions are often most evident in promotion and salary decisions. A familiar litany in many colleges and universities is: "This person is a great teacher, extremely popular with the students, and apparently very effective, but he hasn't written anything and we can't promote him." Similarly, the dilemma faced by the chairperson or personnel committee concerning salary is widely known: "We have limited salary increase funds. "X" has published three papers and a book this year and deserves a major raise, but "Y" has held the undergraduate program together and done a magnificent job with the student group. We do not have enough to give them both what they deserve."

Similar dilemmas intrude when it comes to sanctions. What is to be done about the brilliant researcher who shirks all teaching and committee responsibilities? What about the devoted departmental wheelhorse who can be counted on to serve on any committee but who has abandoned any pretense to scholarship? Such roles cannot be permitted to become typical in a department, but how are the people enacting them to be encouraged to behave differently? This question brings us squarely to issues of power, authority, and decision-making.

Power, Authority, and Decision-Making

It is fairly clear that the bureaucratic organization of colleges and universities impinges more on faculty members and departments on matters of teaching than on matters of research. In his study of 115 universities and colleges in the mid-1960s, Peter Blau (1973) found that teaching activities were more likely to be controlled by bureaucratic forms than were research and scholarly functions. There is evidence that diversity among U.S. universities has increased according to control and prestige. (See Gross and Grambach, 1974.) Of particular importance are the effects of external pressures on the balance between bureaucratic and professional authority. Especially in institutions where professional authority is weak (e.g., community college, church-supported small colleges), external pressures tend to reinforce the bureaucratic structure and to increase centralization of authority.

Some of the pressures from the outside include forces we have already discussed: a general decline in the financial support for higher education, loss of public confidence in colleges and universities, pressures for egalitarian access, and consumer attitudes toward educational products. At the same time, the bargaining power and autonomy of the academic profession have declined: funding for research has levelled off and job mobility has declined drastically.

External pressures on colleges and universities create new roles, for example, affirmative action officers, people whose job is to dispense and monitor federal and state financial aid programs, administrators who oversee human subjects reviews, and others responsible for compliance with federal regulations ranging from occupational safety to facilities for the handicapped. These pressures from the outside have had the effect of increasing the administrative sector.

The organizational context also affects decisions about tenure. As departments fill their quotas for tenured positions, review committees, budget committees, and the chancellor or president take on much more importance than they did during earlier periods of expansion. In response to these forces, administrators put pressure on faculty members. The faculty resist these pressures in the name of autonomy but, as we have seen, the claims to professional expertise by academics rest more on disciplinary scholarship than on teaching. In the name of accountability and productivity --often in response to pressures from legislators, citizens, and other funding sources--central university and college authorities end up speaking for the client, the student. In some schools, administrators rather than faculty become the experts on teaching matters.

Since research and scholarly activities are more insulated from bureaucratic control and teaching activities less insultated, there is a built-in barrier to professional control over teaching; this then increases bureaucratic control over teaching. Those academics in organizations oriented toward research receive the rewards and control the conditions under which they do their scholarly work; they do not need to challenge bureaucratic authority over teaching, since this does not threaten the activities which interest them. Even though more than half of Blau's sample said that teaching ability was extremely important for senior appointments in their institutions--and this fraction is undoubtedly higher now--academics in schools oriented primarily to teaching receive their values and standards from their disciplines. The disciplines until recently have had little to say about teaching as a profession. Thus, increases in centralized controls over teaching are not easily countered by the professionals actually doing the work.

The fundamental organization of higher education in the United States is not likely to change much in the near future. If our analysis has any merit, it should be apparent that the control over teaching--and attention to improving it--has not generally come from the disciplines.

Unions have been more likely to take on some of this role. While the primary goal of a union is economic gain, faculty unions are often concerned as well with non-economic issues, many of them related to teaching. Unions may be a force for control over teaching conditions by the faculty--even a force for improved teaching--as they have tried to be in some places at the elementary and secondary school level.

More initiative should be coming from the disciplines. But as of 1978, few disciplinary associations had a role or structure specifically devoted to dealing with issues of teaching. Few sponsored a journal on teaching, and articles on teaching which had been included in the regular journals in the past have been largely eliminated in recent years. Disciplinary newsletters carry some information on teaching, but most publications sponsored by disciplinary associations are concerned primarily with reporting research and organizational activities.

This situation began to change a few years ago with the formation of an ad hoc committee in the American Sociological Association and the eventual development of a section on teaching. Other disciplinary organizations formed committees on teaching earlier. But even so, there has been almost no communication between disciplinary associations and academic departments--let alone with college administrators--about teaching matters. Individual faculty members, therefore, engage in their teaching essentially unsupported by their own disciplines and often at the mercy of bureaucratic authorities. The various teaching projects of the ASA have reversed this situation for some of the individuals who have been involved in them, thus providing a model for other disciplinary associations. Through these efforts, they may increase faculty members' understanding of and ability to manipulate the points of congruence and incongruence between bureaucratic and disciplinary authority, in the very departments and institutions in which they ply their trade.

References

BLAU, P. (1973) The Organization of Academic Work. New York: Wiley.

GROSS, E. and P. V. GRAMBACH (1974) Changes in University Organization, 1964-1971. New York: McGraw Hill.

Increasing the Reward for Teaching

Reece McGee
Purdue University

This chapter addresses the issue of the reward structure of that portion of the academic profession in major, research-oriented universities in the United States. It is a frequent complaint at universities of this type that one cannot be promoted for teaching, i.e., that organizational reward in the form of promotion, and perhaps to a lesser extent in pay, is positively correlated with research publication and other scholarly activity, and that teaching-related activity goes relatively unrewarded. In its extreme form the complaint may even assert that teaching and related activities are actually punished--that there is a negative correlation between them and reward forthcoming. The problem in short, is the actual activities and consequences of the school's operationalization of the publish-or-perish metaphor. From what I know of the academic profession and its literature, we can probably identify at least three social structural explanations for the phenomenon.

The first of these is simply tradition. There is no tradition in American universities for concern with quality of instruction. This may stem in part from the profession's early adoption of the British collegial model of faculty governance (in the absence of the actual community of scholars represented by the English college). It may also be in part a result of the fact that, in the early years of higher education in the United States, college and, later, university instruction were seen essentially as a finishing process, the object of which was the production of well-rounded gentlemen on the British upper class model. The purpose of education, in other words, was exposure to high culture, which was then presumably absorbed through rote learning and intellectual osmosis.

A second structural explanation is that in the nineteenth century the major universities of the United States adopted the German model of scholarship, a disciplinary model, with the laboratory and speculative investigation at the center of the faculty member's role, rather than the transmission of knowledge to (primarily) undergraduate students. Our universities thus took over a value system in being which defined appropriate role models for faculty members to emulate and research scholarship as the only activity truly appropriate for the professoriate. It is a model which effectively intrudes professional and disciplinary values (i.e., extra-institutional norms) into the organization we call a university. The place where this intrusion is made is the department, the organizational embodiment of the scholarly discipline. Evidence in support of this explanation seems to lie in the fact that the disciplines of the physical sciences were those first imported from Germany, and it is still those fields where the research über alles value system remains most strongly pronounced today.

The third structural explanation for the neglect of (particularly undergraduate) teaching in our universities is the enormous expansion of research opportunities and support after World War II when the federal government and the private foundations poured billions of dollars into the academic research enterprise. This created a new, distinctive, academic workstyle (the word "grantsmanship" was unknown before about 1950) which accentuated the research role at the expense of all other academic obligations. Successful grant procurement became the chief and most honored route to promotion, financial reward, and institutional mobility. The consequence of this situation was the creation of a competitive condition wherein the highest premium was placed on national and international visibility. Individuals and departments had to worry constantly about their scholarly reputations since academic achievement must be constantly renewed in order to maintain reputation. Obsolescence for most scholarly work is quick. In this condition the more parochial concerns of the school, with undergraduate instruction, committee obligations, and so forth, suffer deeply. Organizational service offers no possibility for conversion to reputational credit, no burnishing for individual or departmental prestige in the eyes of others. It is probably not an exaggeration to say that the competition for research recognition at national and international levels actively subverts the possibility of reward for parochial--although essential --activities such as teaching.

Thesis

It is my thesis that as a consequence of these phenomena the major American universities now collectively possess a nationwide, conventionalized, readily accessible model for the evaluation of scholarly work within the several disciplines, a model on which there is high consensus. The degree of consensus seems to vary somewhat by discipline, probably being greatest in what have been called the high paradigm fields, (typically, biophysical sciences), but some significant degree of consensus will be found even in the low paradigm fields of the humanities, with the social sciences lying between the two. Thus, for example, we might expect very high agreement among, say, physicists, that a given piece of work was both (a) appropriately called research, and (b) of a certain quality. Practitioners of, say, English criticism, might have more varied opinions on quality, but would probably be in considerable agreement on the meaning of a work as research (although less, perhaps, than would be found among scientists). The effect of this has been to make the curriculum vitae, or promotion dossier--the actual paper record reviewed by Deans, promotion committees, and department heads when allocating rewards within the university system--a device or record systematically designed to reflect research or scholarly activity in conventionalized ways of a kind permitting inter-subjective agreement. The value system, in short, created the material means and methods for its actualization in reward-granting behavior.

Hypotheses

If the thesis just suggested is accepted, two hypotheses follow. The first is that the existence of the scholarship model, and the institutionalized value system which supports it, inhibit the evolution of other models. It works too well; it is familiar, conventional; we know how to use it. And teaching activity does not fit; the scholarship model is inadequate to measure or demonstrate quality in teaching endeavor, and the value system does not support it as valuable in any event.

The second hypothesis is that, under the present circumstances, (the increase in extra-institutional pressures which urge the universities to respond more centrally to student clientele), the research-oriented value system might be amenable to change if appropriate methods for the evaluation of teaching were invented. It may be, in other words, that in the present situation the universities do not reward teaching more greatly because they do not know how to do so. The only reward-allocating mechanism they possess is unsuited to it, having been invented for the evaluation of other activities. But if an appropriate one were to be made available, its existence might lead to its use.

I need to qualify the second hypothesis in two ways. First, if such a mechanism were invented, it would have to possess the same bureaucratic advantages of the scholarly model. That is, it would have to eventuate in a written record systematically designed to reflect teaching activity in conventionalized ways permitting inter-subjective agreement. It would also have to rest on some kind of generally accepted belief that teaching did have some significance worth rewarding, but that belief, while still in labor in American universities, does seem aborning.

The Case of Purdue - A Kind of Test

My thinking along the lines thus far outlined, and my concern with the problem itself, are in large part an outgrowth of my experience at Purdue University over the past twelve years, and the changes in the university's reward system that have occurred in that period. These changes can be understood as constituting a kind of highly informal and completely uncontrolled test of the hypotheses above. (My own vantage points for observation have been department and school-level promotion committees, three year's service on a Faculty Affairs Committee, and authorship of new school and departmental promotion and tenure regulations and policies.)

The Purdue promotional system is tri-leveled. Departmental committees forward promotion recommendations and supporting documents to school (college) level committees which review them, either rejecting or forwarding them to a university level committee for final action. Promotion regulations require that recommended candidates show a "primary strength" or "excellence" in one of three areas, research, teaching, or public service, and a secondary strength in a second area. In a study of the Humanities and Social Science School's promotional success published during my first year at Purdue, the results revealed that no one had ever been promoted to Full Professor on the primary basis of teaching, and that promotions to Associate Professor on that basis had occurred only in rare instances.

A year or two after this time, the Dean who had headed the school for a number of years retired and was replaced by a much younger, less traditional man. This person determined that one of his goals would be to open up the reward structure to recognition of teaching excellence. To this end, department heads were repeatedly urged to respond to good teaching performances with salary increases, a list of suggestions was circulated to promotion committees indicating kinds of faculty behavior that might at least indicate teaching interest and commitment (e.g., attending teaching workshops, innovation in pedagogical methods, writing about teaching; see Appendix A), and members of promotional committees at both department and school level were repeatedly reminded that teaching was in fact a promotable activity. A related action in part designed to serve these ends was a new policy requiring promotion committee members who cast negative votes on candidates to submit anonymous explanations.

The Dean also took one further, more long-range, action, the fruits of which have yet to be usefully exhibited. A committee was appointed to devise mechanisms for the evaluation of teaching which could be called into action by an individual desiring promotion on those grounds. The committee devoted two or three years of deliberation and effort, including a number of trial runs, before delivering a final report describing a workable procedure (Appendix B). Briefly what is suggested is that, on the behest of the person and in consultation with him, the department head creates a three or four member committee of people both within and without the department with some manifest interest in instruction. These people then meet with the subject, explore his or her teaching goals and philosophies, examine syllabi, classroom handouts, examinations, student evaluations, and so forth, visit classes, interview students, evaluate writings on teaching (if any), and explore such behavior as workshop participation. The committee eventaully produces a written evaluation which is given to the faculty member concerned and to the department head, although the person retains the option of having the report made a part of the promotional document of shown to members of the promotion committee.

It is, as stated, too early to know what effects this system may have on the promotional process at Purdue. It has only been used in one or two cases of actual promotional activity (on one of which I served as an outside member), although several other people have used some variation of it as a device for self-improvement. Many faculty members remain unconvinced of the value of paying this much attention to teaching; the report describing the mechanism was criticized in the school senate for both technical and philosophical shortcomings; some object to the classroom visit provision, and there is considerable wariness that, although the process is procedurally voluntary, if it becomes common it will, in effect, be mandated by promotion committees (as in fact has the use of standardized course-and-instructor evaluation instruments). It is cumbersome; it is immensely time-consuming for committee members; it is not standardized in any scientific sense; and it does not produce results that would be widely accepted as inter-subjectively validatable.

Beyond these matters, there lies an issue more thorny: the procedure does not and cannot measure or demonstrate teaching quality as that term is usually used. The reason, of course, is that we have no conventional agreement on what constitutes good teaching. The procedure does produce a conventional curriculum vitae in which teaching and teaching-related activities may be described in somewhat the same way as research activities and publications, and in which various estimations of teaching quality may also appear. Such a document can demonstrate that a person puts serious effort into teaching, that he or she has communicated such interests to other colleagues, that his or her classroom preparations are imaginative, that students like the person and think well of their experience, and that the teacher is approved by peers as effective and enthusiastic. But it is flawed by the absolute lack of agreement on what constitutes good teaching and the inability to demonstrate such qualities as may be thought to make it up in conventional operationalizations. (Publications, on the other hand, may be read by promotion committees made up of members of the relevant discipline, and persons from other fields may conventionally accept as evidence of quality such things as publication in major disciplinary journals: "I don't know anything about sociology, but if so and so has gotten four pieces into the *American Sociological Review*, he or she must be pretty good.")

We may close this section by noting that, while the particular procedure in question has not been in place long enough to show results, the other urgings and promptings of a powerful and aggressive Dean, and the cooperation of most of the department heads, have resulted in a number of persons being successfully nominated for promotion to associate professor on the primary basis of teaching in the past few years; and last year, one person was promoted to full professor on that basis.

Philosophical, Epistemological and Practical Problems

One of the major elements contributing to the general problem addressed here has been raised above: in the absence of anything even approaching conventional agreement on what constitutes good teaching in a given field, it is extremely difficult to reward people for doing it. One reason for this, although not the only one, is the relative invisibility of the product of instruction. Teaching *activity* is as readily observable as research activity and is, in fact, in one sense *more* public. (It is performed directly before an audience.) But instructional results are not. Research eventuates in some kind of material product: a paper that can be judged by a national or even international audience, a book, an actual artifact, product, or process that others can evaluate or try out for themselves. The results of an instructional experience, to the contrary, are most of them vague, subjective, and perhaps, very long indeed in becoming apparent. One effort to overcome this phenomenon may be represented by the PSI movement, which reduces learning objectives for students to the ability to pass examinations of a standardized variety. Another is the idea of teaching-by-objective which, at least, forces the instructor to articulate what it is he or she is attempting to accomplish and, at best, provides some means of determining whether it has been done. Both of these, of course, remain open to quarrel by those who would argue that simple mastery of technique or subject matter is not what teaching is about, or who would dispute the objectives articulated.

There are, thus, two profound and closely interdependent problems to overcome if we wish to introduce the possibility of systematic professional reward for teaching activity: the problem of definition and the problem of consensus. Until the former is solved, the search for the latter is usually futile. It seems to me hopeless to seek an agreement on what constitutes good or effective teaching; neither morse code nor esperanto would be effective in the Tower of Babel. We might, however, be able to begin to define, and expect reasonably wide agreement on, some list of behaviors essential for minimal satisfaction in instructor performance. Such a list might well be phrased in terms of observable teacher behavior in classroom and office, and of student reactions to it. Barbara Helling's "Guide to Peer Observation" might make a good model with which to begin.[1] (It might be proposed, for example, that it be considered essential to minimal satisfaction in instruction for a faculty member to announce a reasonable number of office hours per week and keep them. Or that some significant proportion of his or her students should agree that he or she usually appears to have prepared for classes beforehand.) If such a list were relatively brief, operationalized where necessary, and eminently reasonable, it is not too much to hope that, if widely circulated, it might achieve some effective consensus. Reasonability seems a critical quality. It would be socially impossible for an opponent to argue in favor of bad teaching, of course, but the possibilities of quibbling are so great, and the opportunities so endless, that for maximum effectiveness, any such device would have to be almost completely unexceptionable, at least upon first introduction. The key epistemological issue is that it would turn our attention away from the fool's errand of seeking a consensual definition of good teaching to the more practical and agreeable objective of defining minimal behavioral competence.

A further bonus appears possible as well. Among the continuing tensions finding their expressive locus in the department is the problem, for sociologists, of the existence of competing paradigms in the field. One of these, already hoary with age, is whether we should consider ours a scholarly discipline or a profession (those with the discipline orientation typically being perceived as holding that research is the only proper goal toward which faculty activity should be addressed). The substitution of minimal teaching competence for excellence should reduce any antagonism toward these efforts from this source. It may be unreasonable to suggest that every sociologist in the country should devote time and effort to becoming an outstanding teacher (indeed, it is impossible), and it might well have serious effect on scholarship in sociology if we were all to do so. But a relatively brief listing of the behavioral configurations of minimal activities essential to basic competence poses no such threat.

Let us examine what some of the desirable characteristics of such a list or procedure might be. I have suggested that we could begin by inventing a kind of checklist of observable behaviors such as promptness in meeting classes, keeping office hours, and so forth. Many of these items could be supplied by student evaluations; others might have to be derived from peers, and certainly the cooperation and participation of the subject would be required. (Some information might be obtainable only by interview, for example.) The object of the exercise would be the production of a document, perhaps a promotion dossier, capable of intersubjective evaluation by others. Such a document would need to show what the person's teaching task is, offer an evaluation of how well and how consistently he or she performs it, and indicate the means by which the judgment was derived. It should certainly reflect the opinions of the student consumers of the teaching as to what they think they got from the experience and how they liked the getting of it (matters on which they are competent to render judgment). And, since college courses are not subjected to editorial review boards composed of subject matter specialists in the way in which journal articles are before they are published, there should be some expression of professional evaluation on the content of what is taught.

Stated in this bald fashion, some of these principles sound controversial indeed, thus contradicting my "reasonableness" criterion offered above. And I am willing to grant that some of the means I would use to effect some of them, if I had the power, would be regarded by many as unreasonable (classroom visits by peers, for example, and full departmental faculty review of syllabi for course content). But we need not demand such implementation. Minimum essential satisfaction can be indicated in other ways, for example an instructor-authored description of course content and objectives, and student course and instructor evaluation forms centering on readily observable, factual matters. The means for doing much of this are already available.

Knowledge Available

There already exists an immense literature on pedagogy. It is widely scattered in the books and journals of a dozen disciplines, some of them unknown to or unrespected by sociologists. A great deal of it is junk and a fair amount of the rest of it is idiosyncratic, inapplicable, or incapable of broad generalization. But interested sociologists and the ASA Projects on Teaching Undergraduate Sociology could pull together what is meritorious and applicable with relative ease. The ASA Projects have already made a good start at this[2] and McKeachie's handbook[3] is a well-known compendium. Two categories of subject have been abundantly researched and can provide solid utility already, and one more gives us some sound information and a lot of possibilities. These are as follows.

Research on student course and instructor evaluation. There is a huge literature on this topic and, in recent years, much of it has begun to converge. I believe it is quite accurate to say that we now know what such paper and pencil evaluations can tell us, under what conditions, and what they cannot. We know that student evaluations are both valid and reliable on some occasions, and not for others, and we know what the conditions are that influence the two. We even know the five general clusters or areas of evaluation involved in what students regard as good teachings, and we know a lot about how student characteristics influence the ratings they give.[4]

Experiments and programs on teaching: The literature of the discipline of education aside (much, but not all of which is irrelevant), there is a small but growing body of knowledge on teaching people--usually graduate students--to teach. It comes from a number of disciplines, and is often convergent. Since a great deal of it is behaviorally oriented, it offers empirical benchmarks for creating a "minimal satisfaction" checklist for instructor behavior in the classroom and office.[5] Neglected resources that might be of considerable value are the very practical, long-standing, and thoroughly tested military Methods of Instruction programs of the armed services, some of which have been going on since World War II.

Research on the relation between teaching and student outcomes: Good research in this area seems relatively rare, but a small body exists which would certainly be useful heuristically; recent work appears to show some signs of convergence.

Knowledge Needed

Much of what we need to know if teaching is to become a rewarded activity in the research-oriented university has already been suggested or implied in preceding pages. A great deal more knowledge is needed on the relations between student outcomes and a variety of student and instructor-related variables. Much of this can only come from careful experimental research, which the attempt to produce the kind of checklist suggested here could itself generate. (We think we know some things about what constitutes satisfactory teaching, and the attempt to produce the kind of device or procedure I have discussed would systematize those beliefs. Very well, then let us test them empirically.)

Unfortunately, we do not know much about altering a collective value system, and that is one of the core problems involved here. But these suggestions imply that perhaps we do not need to expect to change it ourselves. External pressures on the universities are forces for change that the organizations are already responding to in ways that may make ripe the time for the kind of suggestions offered here. In many places it already makes sense to argue that if we do not begin to seek means of making teaching evaluations, others--non-sociologists or even non-academics--will impose them on us, and the preservation of the scholarly aspects of our discipline and its potential utility to our society require that we act now to avert such events. Finally, while there remains considerable entrenched resistance both to the idea of faculty evaluation for teaching, and to the attempt to make teaching a rewardable activity in the university, much of this may be avoidable if we seek merely to assert the possibility of establishing minimal competence rather than excellence as our goal. And as sociologists we do know that even if we cannot guide change, it often occurs by itself, and in foreseeable directions, once institutional mechanisms are provided that make it possible.

Appendix A

TO: Department Heads, School of Humanities, Social Science, and Education

FROM: Robert L. Ringel, Dean

DATE: September 6, 1974

RE: Evaluation of teaching performance with reference to preparation of Promotion Documents

As you know, for some time I have been concerned about documenting "excellence in teaching" insofar as preparing promotion documents are concerned. In one of our recent Council of Department Heads meetings I indicated some categories that might be elaborated upon in preparing promotion documents in which the "teaching" category is to be the one of primary interest. Since that meeting I have continued to read materials pertaining to this issue and would like to take this opportunity to share with you some thoughts concerning categories, areas, criteria, etc. that might pertain to your promotion document preparation. In essence, the list that follows represents some random thoughts on this matter and no intent is made to formulate a comprehensive listing nor to assign any order of priority to the items. I hope that this list will provide you with some ideas that your primary committees and promotion candidates will find acceptable.

Activities of the candidate pertaining to:

a. Advising students

b. Directing student research

c. In-service training for colleagues (on campus, also on national level)

d. Preparation of textbook, tutorial monographs and resource materials

e. Student ratings (quantitative and qualitative); alumni opinions

f. Student recruitment

g. Student placement (marketability of candidate's students)

h. Teaching related committee activity, e.g. graduate committee, curriculum committee, E.P.C., etc.)

i. Activity in teaching research (participation in or development of experimental programs)

j. Community resource activities (consultations)

k. Workshop participation (teacher or student)

l. Activities of self-improvement (courses taken, workshops attended, etc.)

m. Courses developed (candidate's initiative; in response to student's demands)

n. Colleague appraisals (after classroom visits or after studying materials)

o. Awards or recognition of teaching (campus, national)

p. Enrollment pattern in candidate's electives

Appendix B

Report of the Committee on Recognition of Excellence in Teaching (CRET)

This document is the final report of the Committee on Recognition of Excellence in Teaching (CRET), chaired by Professor John F. Feldhusen. The charge that CRET received from my office was to develop a procedure that might serve as an alternate to those procedures currently in use for the evaluation of teaching excellence. The charge was not to develop a procedure which would be the sole method of evaluating teaching nor was it to develop a method which would be a substitute for any or all of the existing procedures. The Committee, in meeting its responsibility, has given us one additional way to look at excellence in teaching.

This document should not be construed as HSSE policy on a method of evaluation of teaching for HSSE has no official policy with regard to this issue.

RLRingel - 2/21/77.

The members of the Committee on Recognition of Excellence in Teaching (CRET) are pleased to present the final revised version of our report. The report represents two and one half years of committee work, tryouts of the procedures, and the use of feedback from the Faculty Affairs Committee, and HSSE Department Heads. Our goal was to develop alternative procedures for the evaluation of faculty who wish to be considered for promotion for excellence in teaching. The Committee does not wish to infringe upon the responsibilities and authority of the departmental primary committees to establish appropriate methods of candidate evaluation.

The Committee recognized that good teaching and scholarship are really inseparable. However, while procedures are available for evaluating scholarship, comparable procedures for the assessment of teaching are lacking. The Committee has attempted to remedy this situation. The report outlines one set of procedures by which objective documentation of teaching abilities may be assembled for the faculty member who chooses to provide such information to promotion committees.

Several other procedures are also available for an assessment of excellence in teaching:

1. The Memorandum from Provost Felix Haas, dated August 22, 1975, stating Policy Concerning Promotions, presents Specifications for Promotion for Excellence in Teaching.
2. Dean Robert L. Ringel's memo to Department Heads, dated September 6, 1974, lists procedures for "Evaluation of teaching performance with reference to preparation of Promotion Documents."
3. Committee C of the American Association of University Professors prepared a statement on procedures for evaluating teaching which was approved by the AAUP Council in June, 1975, and published in

the AAUP Bulletin in August, 1975. (It should be noted that the procedures set forth in this document are in close agreement with the AAUP policy statement.)

4. Finally, the Guidelines for the HSSE Awards for Excellence in Teaching developed by an HSSE committee and approved by Dean Ringel on November 21, 1975, state guidelines for evaluation of excellence in teaching.

Thus, the procedures developed by our Committee and set forth in this document should be viewed as one of a number of alternatives.

Departmental primary committees which elect to use the procedures set forth in this report should remind their faculty in January each year of their availability. Faculty who wish to be evaluated should request that the Department Head, acting as chairman of the primary committee, appoint a committee of three faculty to carry out the evaluation. The committee should be composed of three members who are at least one rank above the candidate. It is suggested that at least one member of the committee be selected from a department other than the candidate's and that this person (or persons) represent a cognate discipline. The Department Head should appoint one of the three selected persons to chair the evaluation committee.

It should be understood that the Department Head acts as chairman of the primary committee in organizing the evaluation committee and that there is mutual agreement between the candidate and the primary committee that the evaluation should be conducted. There would be no point in proceeding with the evaluation if the primary committee felt that it was not an appropriate time in the candidate's career to be considered for promotion for excellence in teaching. It should also be understood that because the time and effort required to conduct an evaluation is extensive, candidates should not expect frequent repetitions of the process if they do not succeed in achieving promotion. A candidate who wishes to be considered for promotion for excellence in teaching should normally request one evaluation for promotion from the assistant to the associate level and one for promotion from associate to full professor. If a candidate is evaluated but fails to be promoted and the primary committee wishes to reconsider the candidate again in a later year, they should reappoint the same committee which conducted the first evaluation to review and update their work and prepare a new report.

If the candidate chooses to use the CRET procedures, the chair of the evaluation committee should meet with the candidate and with the committee to plan the various aspects of the evaluation process using the guidelines which follow. The committee should view these guidelines as suggestions, not prescriptions. There will always be a need to adapt the guidelines and procedures to fit the unique characteristics of the department and discipline. The candidate will also be expected to prepare most of the documentation for the committee.

The evaluation committee members should examine the documents provided by the candidate; should visit at least two classes, one to be selected by the candidate and one by the committee; should interview a randomly selected sample of the candidate's students; and may invite other faculty of the department to provide information concerning the candidate's teaching competence.

The evaluation committee will then prepare a summary report of its fact finding efforts and provide a copy to both the candidate and the Department Head before the end of that Spring semester.

Guide for Self-Evaluation

Evaluation committees will typically expect the candidate to supply information about himself and his teaching as one source of evaluation data. When the committee meets to plan the evaluation, they should specify what information they want from the candidate. The guidelines which follow are suggestive of the information an evaluation committee might find useful. The committee should feel free to select from and to add to this list in order to secure information which they feel contributes to a valid and reliable evaluation of the faculty member's teaching competence.

I. GUIDELINES FOR INITIAL DOCUMENTATION

Courses. Candidates should list and briefly describe the courses they have taught. In the descriptions they may wish to mention such items as enrollments, and whether they designed, or collaborated on the design of, the course(s); whether they redesigned or helped redesign the course(s); and the like.

Impact on Students Here it would be appropriate to have candidates present evidence of their students' (1) achievements on standardized tests, (2) achievements on their own tests or other "teacher-made" tests, (3) preparation for higher level courses, (4) admission to and success in graduate school(s), (5) professional or vocational advancement, and the like. They may also wish to include unsolicited letters of appreciation from former students. In short, they may present any evidence which they regard as important in documenting their impact on students.

Teaching Award(s). Candidates should cite any teaching award(s) they may have received or for which they have been nominated: for example, from a college or university, from a school or division or department therein, or from a professional organization. If they have received no teaching award(s) as yet, this should not prejudice an evaluating team against them.

Growth as a Teacher. Candidates should cite colloquia or workshops on teaching, which they have attended and/or the contributions(s) they made or the role(s) they filled. They may also wish to list the state and national professional organizations they support by means of membership or in any other ways. For example, on one or another of these organizations they may be serving as a member or officer of a committee devoted to curriculum improvement, teacher preparation, or the like.

Instructional Materials and Approaches. If candidates have developed innovative instructional materials and approaches they should briefly describe them and cite any external and/or internal grant(s) they may have received for this purpose. In addition--or as an alternative--they should briefly describe how they have "field-tested" another teacher's innovations with a view to adopting or adapting them. What criteria have they used for evaluating their own materials or those of a colleague in their field? Title(s), publisher(s), and date(s) of publication of one or more textbooks or other instructional material(s) they have written or collaborated on should also be cited. (NOTE: For collaborations they should note the extent--e.g., equal, major, minor--of their participation.) They should also list title(s) of their speech(es) and paper(s), citing date(s), place(s), etc., and whether published or accepted for publication, again citing appropriate bibliographical details.

Other Teaching Role(s). They should briefly describe the contribution(s) they have made in one or another of such situations as (1) working with graduate students--for example, in helping place graduates, etc.; (2) supervising students and/or student-teachers in field or clinic experiences; (3) training, "mentoring," or coordinating teaching assistants; (4) giving workshops for practitioners in the field; (5) tutoring or arranging for tutoring of students; (6) serving as faculty fellow or adviser to student organizations; (7) arranging internships; and (8) any other role not mentioned among the examples above.

II. DESCRIBING AND DOCUMENTING TEACHING STYLE(S)

A. Teaching Style(s). Teaching spans a wide spectrum of styles. Evaluating teams will be searching for excellence in teaching wherever it appears. Therefore, candidates should indicate the principal strategy (or strategies) they use in their teaching. They may wish to nominate, for example, one or more of the following modes of teaching: (1) expository --i.e., the transmission of knowledge primarily by lecture; (2) discussive--i.e., Socratic questioning, structured lecture-discussions, or unstructured approaches; (3) team teaching; (4) working with small groups; (5) individual projects or any other kinds of individualizing if the principal strategy is "individualizing," see item B.3 below; (6) laboratory experiences; (7) case methods; (8) any other strategy or mode not mentioned among the possibilities above.

B. Documentation. To illustrate the principal strategy or strategies nominated above, candidates should document each of the following four subcategories wherever applicable.

1. Course Goals and Objectives. They should state their goals and objectives in terms of (a) what learning is expected, (b) the criteria for evaluating it, and (c) the relative weights accorded the stated objectives.

2. Course Materials. They should present such items as a course outline or syllabus; representative assignment(s); handouts--including study guides, bibliographies, etc.; copies of materials deposited at a study center or the Reserve Book Room; current textbook(s) or bibliographical identification(s) thereof; audio-visual media and materials--e.g., film titles and tapes; representative quizzes and examinations; and the like. They might well bear in mind that evaluators of the instructional materials (e.g. the syllabus and text titles) look for evidence of up-to-date scholarship and its application(s).

3 Individualization of Instruction. If applicable, they may wish to add descriptions of such special teaching strategies as autotutorial units, PSI courses, mini-courses, learning centers, office

hours, and any other means devised and put into practice to accommodate the needs and interests of individual students.

4. Evaluation of Teaching. Here they would ordinarily include not only their own evaluation(s)--e.g., those based on their use of tape recordings and/or videotapes--but also evaluations elicited from colleagues and students.

GUIDELINE QUESTIONS FOR EVALUATORS

In carrying out the evaluation and in preparing the evaluation report, the committee should attempt to organize its efforts according to some guidelines or criteria. The following questions are suggested as possible guidelines. The committee might also develop other guidelines appropriate for the candidate and the department.

I. Goals and/or Objectives. Are they presented to the students? Do they seem appropriate for the course? Do they correlate with instructional activities? Are they related to the tests and to grading? Are the weights or emphases assigned to the various components of the overall grade explained to the students at the outset of the course? Are goals and objectives used as motivators? Are students helped to set their own goals?

II. Materials for Course. Is a course outline or syllabus given to students? Is the presented bibliography up-to-date and reflective of recent scholarship and knowledge of the field? Are the materials that are deposited at a study center reflective of excellence in selection or development? Are the assignments appropriate for the course and goals? Are they evaluated promptly and reliably?

III. Audio-visual Media and Material. Are a variety of media used? Are they used effectively? Do they fit course goals and objectives?

IV. Communication of Knowledge.

Expository: Is the instructor skillful in lecturing? Does the instructor explain difficult concepts well? Are good illustrations used? Does the instructor summarize well? Does the instructor ask questions in such a fashion as to stimulate thought?

Discussion: Does the instructor use small groups effectively in teaching? If an unstructured approach is used, does the instructor act effectively as a catalyst? If the instructor uses a Socratic approach, are major points developed skillfully through questioning?

Team Teaching: Does the instructor work well with others in designing the course? In teaching the course? Does the instructor coordinate teaching team?

Individualization of Instruction: Are mini-courses, auto-tutorial units, projects, etc. designed to fit individual interests and needs?

Laboratory. Are sessions integrated with other course activities? Are sessions well planned and carried out?

Reinforcement and Encouragement: Are students encouraged by the instructor and reinforced for their efforts to learn?

Humor: Is humor used to motivate students and to communicate knowledge effectively?

Simulations and Games: Does the instructor use simulations and games to enhance learning?

Demonstrations and Illustrations: Is the course made interesting through skillful and appropriate use of examples, illustrations, and demonstrations?

Variety of Techniques: Does the instructor use a variety of approaches to help students learn?

V. Testing, Evaluation and Grading: Do the tests fit the course goals and objectives? Are the testing and grading policies clearly stated? Are test results returned promptly? Does the instructor acquaint the students with the nature of exams in advance and critique results?

VI. Quality of Scholarship for Teaching: Do course materials reflect good scholarship? What is the quality of scholarship as reflected in texts and reading assignments selected for courses?

VII. Overall Organization and Administration of Course: Is the course well planned and organized? Is it managed well? Is it flexible enough to meet student needs? How well does the instructor carry out the principal teaching style nominated in II. A. (page 4)?

VIII. Interactions Outside of Class: Are students helped through office contacts? Are special help or test preparation sessions conducted? Does the instructor critique tests with individual students? Does the instructor meet with small groups of students?

IX. Personal Characteristics: Does the instructor demonstrate enthusiasm for teaching? Does the instructor show concern for individual students? Is the instructor tactful and courteous to students? Does the instructor serve as a scholarly model for students? Is the instructor friendly and considerate? Does the instructor respect students and develop rapport with them? Is the instructor flexible and willing to change course procedures? Does the instructor offer to assist students outside of class?

X. Course and Teaching Evaluation: Is some evaluation system used? Are there indications that courses and/or instructional styles have been altered by past evaluations?

XI. Other:

Notes

1. Barbara Helling, "Looking for Good Teaching: A Guide to Peer Observation," Danforth Faculty Fellowship Project Report, St. Olaf College, Northfield, Minnesota. 1976.

2. Beginners can start with "Teaching Resources Available," Teaching Resource Center, ASA Executive Office, Washington, D.C. 20036.

3. Wilbert J. McKeachie, Teaching Tips: A Guidebook for the Beginning College Teacher (Lexington, Mass.: D. C. Heath & Co., 7th ed. 1978).

4. The areas, as summarized by Ted Wagenaar of Miami University, are (1) skill/communication; (2) structure/organization; (3) dynamism/enthusiasm/rapport; (4) difficulty/workload; (5) knowledge of subject. See also Gary D. Borich and Susan K. Miller, Evaluating Classroom Instruction: A Sourcebook of Instruments. (Reading, Mass.: Addison-Wesley Publishing Co., 1977), Kenneth P. Doyle, Student Evaluation of Instruction (Lexington, Mass.: Lexington Books, 1975), Richard Miller, Evaluating Faculty Performance (San Francisco: Jossey-Bass, 1972).

5. See, for example, Charles Goldsmid and Everett K. Wilson, forthcoming book on teaching sociology. (Belmont, California: Wadsworth Publishing Co., 1980).